

```
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1113
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum
US-10-284-138-3

Query Match          21.0%; Score 57; DB 15; Length 1113;
Best local Similarity 31.4%; Pred. No. 1.4e+02;
Matches 16; Conservative 8; Mismatches 27; Indels 0; Gaps 0;

QY      1 QLNHEGFKLFATEATSDWLNANNVPATPVAWPSQEGQNPSLSIRKLIRDG 51
       :|: |:|: ||| |: | |: | |: | |: | |: | |: | |: | |: |
Db      994 RLASMGYKILATEGTAGMLRRNGIDCEVVLKASDIREGVGEKSIVDRIREG 1044

Search completed: September 13, 2003, 02:28:05
Job time : 38 secs
```

```
Query Match      22.1%; Score 60; DB 15; Length 70;
Best Local Similarity 52.4%; Pred. No. 2.3;
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY      31 WPSQEGQNPSLSIRKLIRDG 51
      :|||:||||:|:|:|
Db      39 YPGQGETPSLKIQLVRCG 59

RESULT 10
US-09-468-147-91
; Sequence 91, Application US/09468147A
; Publication No. US20030049601A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232.US.P1
; CURRENT APPLICATION NUMBER: US/09/468,147A
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/173,141
; EARLIER FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: US 60/061,199
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 91
; LENGTH: 1698
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: xaa = Unknown or Other at position 174
; OTHER INFORMATION: xaa = Unknown or Other at position 363
; OTHER INFORMATION: xaa = Unknown or Other at position 1088
; OTHER INFORMATION: xaa = Unknown or Other at position 1131
; OTHER INFORMATION: xaa = Unknown or Other at position 1217
; OTHER INFORMATION: xaa = Unknown or Other at position 1389
US-09-468-147-91

Query Match      22.0%; Score 59.5; DB 11; Length 1698;
Best Local Similarity 44.4%; Pred. No. 1.1e+02;
Matches 12; Conservative 4; Mismatches 10; Indels 1; Gaps 1;

QY      3 HNEGFKLFATEATSDWLNANNVPATPV 29
      | | | : | | : | | | | |
Db      911 HRPGEDELYLTPAANWFEANK-PAQPV 936

RESULT 11
US-09-285-385C-4
; Sequence 4, Application US/09285385C
; Patent No. US20020150946A1
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S.
; APPLICANT: Scott, Ian C.
; APPLICANT: Thomas, Christina L.
; TITLE OF INVENTION: MAMMALIAN TOLLOID-LIKE GENE AND PROTEIN
; FILE REFERENCE: 960296.96111
; CURRENT APPLICATION NUMBER: US/09/285,385C
; CURRENT FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/111873
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/080550
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 4
; LENGTH: 1012
```

```
; TYPE: PRT
; ORGANISM: mouse
US-09-285-385C-4

Query Match      21.8%; Score 59; DB 10; Length 1012;
Best Local Similarity 25.4%; Pred. No. 69;
Matches 15; Conservative 7; Mismatches 25; Indels 12; Gaps 2;

QY      2 LHNEGFKLFATEATSDWL-----NANNVPATPVANPQEQGNPSSIRKLIRDG 51
      | : ||| : || : | | | : | | : | | : | |
Db      65 LDEDDLKLFHIDKAEDWTKPSIDKPGHDTGGLLETSARWPNDTASNASIQAPRK---DG 120

RESULT 12
US-10-029-180-102
; Sequence 102, Application US/10029180
; Publication No. US20020182708A1
; GENERAL INFORMATION:
; APPLICANT: Cali, Brian M.
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin T.
; APPLICANT: Milna, G. Todd
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeffrey C.
; APPLICANT: Trueheart, Josh
; APPLICANT: Zhang, Lixin
; TITLE OF INVENTION: No. US20020182708A1e1 Regulators of Fungal Gene Expression
; FILE REFERENCE: MIC-004
; CURRENT APPLICATION NUMBER: US/10/029,180
; CURRENT FILING DATE: 2001-12-22
; PRIOR APPLICATION NUMBER: US 60/257,431
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 102
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fungal gene
US-10-029-180-102

Query Match      21.2%; Score 57.5; DB 14; Length 576;
Best Local Similarity 44.8%; Pred. No. 56;
Matches 13; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

QY      15 TSDWLNANNVPATPVA-WPSQEQGNPSSL 42
      | | | | | | | | | | | | |
Db      96 TPDWSLQNTVPFMPVATPFPSSSSTSPGLT 124

RESULT 13
US-09-494-359-3
; Sequence 3, Application US/09494359
; Publication No. US20030124685A1
; GENERAL INFORMATION:
; APPLICANT: Kuwabara, Yoko
; APPLICANT: HASHIGUCHI, Kenichi
; APPLICANT: NAKAMATSU, Tsuyoshi
; APPLICANT: KURAHASHI, Osamu
; APPLICANT: MORI, Yukiko
; APPLICANT: ITO, Hisao
; TITLE OF INVENTION: CARBAMOYL-PHOSPHATE SYNTHETASE GENE OF CORNEIFORM BACTERIA A
; TITLE OF INVENTION: PRODUCING L-ARGININE
; FILE REFERENCE: 0010-1082-0
; CURRENT APPLICATION NUMBER: US/09/494,359
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: JP 11-24149
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1018
```

;
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
;

;
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
;

QY **2** LHNEGKLFATEATSDWLNANNVPATPV--AWPSQEQGNPSSLSSIR 45
 :
 :
D6 **28** VYNHAFRL-----TADWSTAEDWMATTFLAEWLLRERYDPGGSLR 68

```

RESULT 6
US-09-764-878-129
; Sequence 129, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA121
; CURRENT APPLICATION NUMBER: US/09/764,878
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 129
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-878-129

```

Query Match	22.1%;	Score 60;	DB 9;	Length 70;
Best Local Similarity	52.4%;	Pred. No. 2.3;		
Matches 11;	Conservative	4;	Mismatches 6;	Indels 0;
			Gaps 0;	

```

QY      31 WPSQEGQNPSSLIRKLIRDG 51
      :| | |: ||| |:|:| |
Db      39 YPGQHGETPSLLKIQKLVRCG 59

```

```

RESULT 7
US-09-764-860-464
; Sequence 464, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 464
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-860-464

```

Query Match	22.1%	Score 60;	DB 9;	Length 70;
Best Local Similarity	52.4%;	Pred. No. 2.3;		
Matches 11; Conservative	4;	Mismatches 6;	Indels 0;	Gaps 0;

```

Qy      31 WPSQEGQNPSSLSSIRKLIRDG 51
          :| | | :| | | | :| | | |
Db      39 YPGQHGETPSLLKIOKLVRCG 59

```

RESULT 8
US-10-079-854-129
; Sequence 129, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA121C1
; CURRENT APPLICATION NUMBER: US/10/079,854
; CURRENT FILING DATE: 2002-02-22
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 428

```

; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 129
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-079-854-129

```

Query Match	22.1%	Score 60;	DB 15;	Length 70;
Best Local Similarity	52.4%	Pred. No. 2.3;		
Matches 11:	Conservative	4:	Mismatches 6;	Indels 0;
				Gaps 0;

```

QY      31 WPSQEGQNPSLSSIRKLIRDG 51
          :| | |: | | | |: | | |
Db      39 YPGOHGETPSSLKIOKLVRCG 59

```

RESULT 9

US-10-074-095-464

; Sequence 464, Application US/10074095

; Publication No. US20030077704A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC008C1

; CURRENT APPLICATION NUMBER: US/10/074,095

; CURRENT FILING DATE: 2002-02-14

; PRIOR APPLICATION NUMBER: 09/764,860

; PRIOR FILING DATE: 2001-01-17

; PRIOR APPLICATION NUMBER: 60/179,065

; PRIOR FILING DATE: 2000-01-31

; PRIOR APPLICATION NUMBER: 60/180,628

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: 60/214,886

; PRIOR FILING DATE: 2000-06-28

; PRIOR APPLICATION NUMBER: 60/217,487

; PRIOR FILING DATE: 2000-07-11

; PRIOR APPLICATION NUMBER: 60/225,758

; PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/220,963

; PRIOR FILING DATE: 2000-07-26

; PRIOR APPLICATION NUMBER: 60/217,496

; PRIOR FILING DATE: 2000-07-11

; PRIOR APPLICATION NUMBER: 60/225,447

; PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/218,290

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: 60/225,757

; PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/226,868

; PRIOR FILING DATE: 2000-08-22

; PRIOR APPLICATION NUMBER: 60/216,647

; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: 60/225,267

; PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/216,880

; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: 60/225,270

; PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/251,869

; PRIOR FILING DATE: 2000-12-08

; PRIOR APPLICATION NUMBER: 60/235,834

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: 60/234,274

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: 60/234,223

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: 60/228,924

; PRIOR FILING DATE: 2000-08-30

; PRIOR APPLICATION NUMBER: 60/224,518

; PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/236,369

; PRIOR FILING DATE: 2000-09-29

; PRIOR APPLICATION NUMBER: 60/224,519

Sequence 5059, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5059
LENGTH: 344
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5059
Query Match 25.6%; Score 69.5; DB 9; Length 344;
Best Local Similarity 37.1%; Pred. No. 0.89;
Matches 13; Conservative 7; Mismatches 14; Indels 1; Gaps 1;
QY 1 QLNHEGFKLFATEATSDWL-NANNVPATPVAPSQ 34
Db 61 RVHSGFVRFLQNAWDWLATGRSHDMLPIAWPTR 95
RESULT 3
US-10-169-048-40
Sequence 40, Application US/10169048
Publication No. US20030072769A1
GENERAL INFORMATION:
APPLICANT: Clarke, Edna Elizabeth
APPLICANT: Zhou, Liqing
APPLICANT: Shea, Jacqueline Elizabeth
APPLICANT: Feldman, Robert Graham
APPLICANT: Holden, David William
TITLE OF INVENTION: Streptococcus Pyogenes Virulence Genes and Proteins And Their Use
FILE REFERENCE: GJE-97
CURRENT APPLICATION NUMBER: US/10/169,048
CURRENT FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: PCT/GB00/04997
PRIOR FILING DATE: 2000-12-22
NUMBER OF SEQ ID NOS: 62
SOFTWARE: PatentIn version 3.1
SEQ ID NO 40
LENGTH: 339
TYPE: PRT
ORGANISM: Streptococcus pyogenes
US-10-169-048-40
Query Match 23.6%; Score 64; DB 15; Length 339;
Best Local Similarity 38.8%; Pred. No. 4.4;

Matches 19; Conservative 7; Mismatches 19; Indels 4; Gaps 2;
QY 2 LHNKGFKLFATEATSDWLNANNVPATPVAPSQEQNP-SLSIRKLIR 49
Db 174 LSNEGFAIFAGTAVND--TSSVTATATAWDALHHSNTLDGATIVKLTR 219
RESULT 4
US-09-976-059-30
Sequence 30, Application US/09976059
Patent No. US20020164747A1
GENERAL INFORMATION:
APPLICANT: Farnet, Chris
APPLICANT: Zazopoulos, Emmanuel
APPLICANT: Staffa, Alfredo
TITLE OF INVENTION: Genes and Proteins for Biosynthesis of Ramoplanin
FILE REFERENCE: 3019-PCT
CURRENT APPLICATION NUMBER: US/09/976,059
CURRENT FILING DATE: 2001-10-15
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn version 3.0
SEQ ID NO 30
LENGTH: 619
TYPE: PRT
ORGANISM: Actinoplanes sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1)
OTHER INFORMATION: V represents a non-standard initiator codon. It is expected t
OTHER INFORMATION: the biosynthesized protein will have a formylmethionine residu
OTHER INFORMATION: at this position
US-09-976-059-30
Query Match 22.9%; Score 62; DB 10; Length 619;
Best Local Similarity 34.5%; Pred. No. 16;
Matches 19; Conservative 7; Mismatches 21; Indels 8; Gaps 3;
QY 3 HNEGFKLFATEATSDWLNANN--VPA---TPVAPSQEQNPSSLIRKLIRDG 51
Db 514 HRDGRTYLA--ATSSWRTAGRLIVPTGHSYLPGLGFSGAAPFPSLAGVQRLVRDG 566
RESULT 5
US-10-156-761-11422
Sequence 11422, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 11422
LENGTH: 232
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-11422
Query Match 22.3%; Score 60.5; DB 15; Length 232;
Best Local Similarity 34.8%; Pred. No. 8;
Matches 16; Conservative 8; Mismatches 15; Indels 7; Gaps 2;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2003, 02:16:53 ; Search time 36 Seconds
(without alignments)
205.709 Million cell updates/sec

Title: US-09-585-077C-4_COPY_1380_1430
Perfect score: 271
Sequence: 1 QLNHEGFKLFATEATSDWLN.....PSQEQNPSSLIRKLIRDG 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 541936 seqs, 145912426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*

10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*

17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	71	26.2	1102	15	US-10-156-761-14395
2	69.5	25.6	344	9	US-09-815-242-5059
3	64	23.6	339	15	US-10-169-048-40
4	62	22.9	619	10	US-09-976-059-30
5	60.5	22.3	232	15	US-10-156-761-11422
6	60	22.1	70	9	US-09-764-878-129
7	60	22.1	70	9	US-09-764-860-464
8	60	22.1	70	15	US-10-079-854-129
9	60	22.1	70	15	US-10-074-095-464
10	59.5	22.0	1698	11	US-09-468-147-91
11	59	21.8	1012	10	US-09-285-385C-4
12	57.5	21.2	576	14	US-10-029-180-102
13	57	21.0	1018	11	US-09-494-359-3
14	57	21.0	1113	9	US-09-836-470B-3
15	57	21.0	1113	15	US-10-284-138-3

16	57	21.0	1113	15	US-10-284-334-3	Sequence 3, Appli
17	56.5	20.8	504	15	US-10-156-761-12417	Sequence 12417, A
18	56.5	20.8	673	15	US-10-157-031-291	Sequence 291, App
19	55.5	20.5	823	15	US-10-177-293-240	Sequence 240, App
20	55.5	20.5	1435	15	US-10-128-714-8125	Sequence 8125, Ap
21	55	20.3	441	9	US-09-879-957-34	Sequence 34, Appli
22	55	20.3	1113	10	US-09-738-626-5279	Sequence 5279, Ap
23	54.5	20.1	1228	10	US-09-117-447-2	Sequence 2, Appli
24	54	19.9	100	10	US-09-872-523-8	Sequence 8, Appli
25	54	19.9	101	15	US-10-097-065-576	Sequence 576, App
26	54	19.9	140	12	US-10-238-075-984	Sequence 984, App
27	54	19.9	570	9	US-09-815-242-10760	Sequence 10760, A
28	54	19.9	3739	10	US-09-861-289-33	Sequence 33, Appli
29	54	19.9	3739	10	US-09-860-846-33	Sequence 33, Appli
30	54	19.9	3739	11	US-09-988-384B-33	Sequence 33, Appli
31	54	19.9	3739	11	US-09-836-821-33	Sequence 33, Appli
32	54	19.9	3739	11	US-09-793-708-2	Sequence 2, Appli
33	54	19.9	3739	12	US-10-201-365-3	Sequence 3, Appli
34	54	19.9	3739	12	US-10-160-539-2	Sequence 2, Appli
35	54	19.9	11877	10	US-09-861-289-6	Sequence 6, Appli
36	54	19.9	11877	10	US-09-860-846-6	Sequence 6, Appli
37	54	19.9	11877	11	US-09-836-821-6	Sequence 6, Appli
38	54	19.9	12199	11	US-09-988-384B-6	Sequence 6, Appli
39	53.5	19.7	311	15	US-10-156-761-14936	Sequence 14936, A
40	53.5	19.7	639	10	US-09-906-209-2	Sequence 2, Appli
41	53	19.6	276	14	US-10-033-078-16	Sequence 16, Appli
42	53	19.6	288	15	US-10-205-823-194	Sequence 194, App
43	53	19.6	401	9	US-09-062-113-63	Sequence 63, Appli
44	53	19.6	401	12	US-10-232-858-63	Sequence 63, Appli
45	53	19.6	452	12	US-10-238-075-859	Sequence 859, App

ALIGNMENTS

RESULT 1

US-10-156-761-14395

; Sequence 14395, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 14395

; LENGTH: 1102

; TYPE: PRT

; ORGANISM: Streptomyces avermitilis

US-10-156-761-14395

Query Match 26.2%; Score 71; DB 15; Length 1102;
Best Local Similarity 40.4%; Pred. No. 2.2;
Matches 21; Conservative 8; Mismatches 21; Indels 2; Gaps 2;

Qy 1 QLNHEGFKLFATEATSDWLNANNVPATPVAMPSEQQNP-SLSSIRKLIRDG 51
Db 981 ELVAHGFEILLATSGAEVLKRNGINATVVRKQS-EGEGPGGEKTIVQLIHDG 1031

RESULT 2
US-09-815-242-5059


```
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/723,820
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/541,782
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: S.pombe
US-09-723-820-6

Query Match      19.9%; Score 54; DB 4; Length 1073;
Best Local Similarity 31.4%; Pred. No. 1.6e+02;
Matches 16; Conservative 8; Mismatches 21; Indels 5; Gaps 2;

QY      5  EGFKLFAEATSDWLNANNVPATPVAPPS-----QEGQNPSLSSIRKLIRD 50
      | | | | | : ||||| : | | | | | : | | | | | :
Db      986  EDTSLVKLETTGDTPTPSKRELEATP-SWTRDSSLIKETTNNLNLDSDKKFVRE 1035
```

Search completed: September 13, 2003, 02:17:36
Job time : 22 secs

```

; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-630-915A-34

Query Match 20.3%; Score 55; DB 4; Length 441;
Best Local Similarity 37.5%; Pred.No. 38;
Matches 9; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 18 WLNANNVPATPVAVPSQEGQNPSL 41
| : | | | | | | | | |
Db 283 WFRSTSGPGMPMNPQPEEWNPD 306

RESULT 13
US-09-107-532A-6689
; Sequence 6689, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6689:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 569 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...569
; SEQUENCE DESCRIPTION: SEQ ID NO: 6689:
US-09-107-532A-6689

Query Match 19.9%; Score 54; DB 4; Length 569;
Best Local Similarity 23.9%; Pred.No. 71;
Matches 11; Conservative 12; Mismatches 23; Indels 0; Gaps 0;

QY 2 LHNEGKLFATEATSDWLNANNVPATPVAVPSQEGQNPSLSIRKL 47
| | : : | | | | | | | | : : : : |
Db 466 LQNDQLTMKTAKTDTYWLHAKNIPGSHVVIKSKDKPSDETITEAEL 511

RESULT 14
US-09-541-782-6
; Sequence 6, Application US/09541782
; Patent No. 6284480
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/541,782
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: S.pombe
; US-09-541-782-6

Query Match 19.9%; Score 54; DB 3; Length 1073;
Best Local Similarity 31.4%; Pred.No. 1.6e+02;
Matches 16; Conservative 8; Mismatches 21; Indels 6; Gaps 2;

QY 5 EGFKLFATEATSDWLNANNVPATPVAVPS-----QEGQNPSLSIRKL 50
| | | | | : : | | | | : : | : |
Db 986 EDTSLVKLETTGDTPSKRELPA TP-SWTRDSSLIKETTNNLSDSKKFVRE 1035

RESULT 15
US-09-723-820-6
; Sequence 6, Application US/09723820
; Patent No. 6468760
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey

```



```

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27544
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27544

Query Match      25.6%; Score 69.5; DB 4; Length 349;
Best Local Similarity 37.1%; Pred. No. 0.35;
Matches 13; Conservative 7; Mismatches 14; Indels 1; Gaps 1;

QY      1 QLNHEGFKLFATPATSDWL-NANNVPATPVAWPSQ 34
      ::::| | | | | : | | | | : | | | | :
Db      66 RVHSEGEVRFQLQNAQDWLATGRSHDMLPIAWPTR 100

RESULT 6
US-09-252-991A-27861
; Sequence 27861, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27861
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27861

Query Match      23.2%; Score 63; DB 4; Length 96;
Best Local Similarity 43.8%; Pred. No. 0.49;
Matches 14; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY      14 ATSDWLNANNVPATPVAWPSQEGQNPSLSSIR 45
      :::| | | | | | | | | | | | | | |
Db      47 SSSPWPATRPATSVANTSPACTRPPCSRYR 78

RESULT 7
US-09-134-001C-4694
; Sequence 4694, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4694
; LENGTH: 228

```


Q9ESY1
ID Q9ESY1 PRELIMINARY; PRT; 1803 AA.
AC Q9ESY1;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE TEMO.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Mruk D., Mo M.-Y., Cheng C.Y.;
RT "TMO is a marker to study sertoli-germ cell interactions: cloning and
RT regulation of a novel testicular molecule.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF296169; AAG09060.1; -
SQ SEQUENCE 1803 AA; 206588 MW; ED73B36E135F25B2 CRC64;

Query Match 23.4%; Score 63.5; DB 11; Length 1803;
Best Local Similarity 37.8%; Pred. No. 62;
Matches 14; Conservative 7; Mismatches 9; Indels 7; Gaps 2;

QY 1 QLNHEGFKLFATEATSDWLNANNVPATPVAVWPSEQG 37
Db 185 ELHHKGFKLWFEDELIGLVSVQNLP----QW---EGQ 214

RESULT 15
Q9N973
ID Q9N973 PRELIMINARY; PRT; 2921 AA.
AC Q9N973;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical 303.3 kDa protein.
GN P1295.15.
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Masuy D., Purnelle B., Goffeau A., Ivens A.C., Quail M.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL; AL359773; CAB95223.1; -
DR InterPro; IPR000183; Decarboxylase2.
DR PROSITE; PS00879; ODR_DC_2_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 2921 AA; 303327 MW; FD6B51F33AE37470 CRC64;

Query Match 23.4%; Score 63.5; DB 5; Length 2921;
Best Local Similarity 34.0%; Pred. No. 1.1e+02;
Matches 16; Conservative 10; Mismatches 20; Indels 1; Gaps 1;

QY 4 NEGFKLFATEATSDWLNANNVPATPVAVWPSEQGNPSLSIRKLIRD 50
Db 2277 NDGEGLLRTRSDADSRARRV-STPTPTPPPPQPPQPSGASLQRTLAD 2322

Db 174 LSNEGFAIFAGTAVND--TSSVTATATAWDALHHSNTLDGATIVKLT 219

RESULT 11

Q8K7M5 PRELIMINARY; PRT; 339 AA.

AC Q8K7M5;

DT 01-OCT-2002 (TReMBLrel. 22, Created)

DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)

DE Hypothetical protein Spym3_0740.

GN SPYM3_0740.

OS Streptococcus pyogenes (serotype M3).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=198466;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MGAS315 / Serotype M3;

RX MEDLINE=22133808; PubMed=12122206;

RA Beres S.B., Sylva G.L., Barbian K.D., Lei B., Hoff J.S.,

RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,

RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,

RA Schlievert P.M., Musser J.M.;

RT "Genome sequence of a serotype M3 strain of group A Streptococcus:

RT phage-encoded toxins, the high-virulence phenotype, and clone

RT emergence."

RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).

DR EMBL; AE014151; AAM79347.1; -

DR InterPro; IPR003439; ABC_transporter.

DR InterPro; IPR004630; Cons_hypoth698.

DR Pfam; PF03601; Cons_hypoth698; 1.

DR PROSITE; PS00211; ABC_TRANSPORTER; 1.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 339 AA; 36275 MW; 4C1D2C9A3C94283B CRC64;

Query Match 23.6%; Score 64; DB 16; Length 339;

Best Local Similarity 38.8%; Pred. No. 8;

Matches 19; Conservative 7; Mismatches 19; Indels 4; Gaps 2;

QY 2 LHNEGFKLFATEATSDWLNANNVPATPVAMPVPSQEQGNP-SLSIRKILR 49

Db 174 LSNEGFAIFAGTAVND--TSSVTATATAWDALHHSNTLDGATIVKLT 219

RESULT 12

Q8UG58 PRELIMINARY; PRT; 863 AA.

AC Q8UG58;

DT 01-JUN-2002 (TReMBLrel. 21, Created)

DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)

DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)

DE DNA-directed RNA polymerase.

GN ATU1183 OR AGR_C_2186.

OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Rhizobium.

OX NCBI_TaxID=176299;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21608550; PubMed=11743193;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,

RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,

RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,

RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,

RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,

RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,

RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,

RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,

RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,

RA Nester E.W.;

RT "The genome of the natural genetic engineer Agrobacterium tumefaciens

RT C58.";

RL Science 294:2317-2323(2001).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=21608551; PubMed=11743194;

RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin L.,

RA Qurolo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Liu F.,

RA Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,

RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,

RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,

RA Cielo C., Slater S.;

RT "Genome sequence of the plant pathogen and biotechnology agent

RT Agrobacterium tumefaciens C58.";

RL Science 294:2323-2328(2001).

DR EMBL; AE009081; AAL42195.1; -

DR EMBL; AE008047; AAK86987.1; -

DR InterPro; IPR002092; RNA_pol_phage.

DR Pfam; PF00940; RNA_pol; 1.

DR PROSITE; PS00900; RNA_POL_PHAGE_1; 1.

DR PROSITE; PS00489; RNA_POL_PHAGE_2; 1.

KW DNA-directed RNA polymerase; Complete proteome.

SQ SEQUENCE 863 AA; 97075 MW; 9B038C4FB44D36CA CRC64;

Query Match 23.6%; Score 64; DB 16; Length 863;

Best Local Similarity 30.2%; Pred. No. 23;

Matches 19; Conservative 7; Mismatches 23; Indels 14; Gaps 2;

QY 3 HNEGFKLFATEATSDWL--NANNVPATPVAMPVPSQEQGNPSSLSSIR-----KLI 48

Db 221 HNTNIYVATKETLDWLATENSRLAPLSPVLTLPVPPRWTSPFRGGYWSGRVRLRI 280

QY 49 RDG 51

Db 281 KTG 283

RESULT 13

Q14997 PRELIMINARY; PRT; 1798 AA.

AC Q14997;

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)

DE Hypothetical protein KIAA0077 (Fragment).

GN KIAA0077.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96051398; PubMed=7584044;

RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,

RA Seki N., Kawarayashi Y., Ishikawa K., Tabata S.;

RT "Prediction of the coding sequences of unidentified human genes. II.

RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by

RT analysis of cDNA clones from human cell line KG-1.";

RL DNA Res. 1:223-229(1994).

DR EMBL; D38521; BAA07526.1; -

KW Hypothetical protein.

FT NON_TER 1

SQ SEQUENCE 1798 AA; 206406 MW; 6FCB2F2AA77999C1 CRC64;

Query Match 23.4%; Score 63.5; DB 4; Length 1798;

Best Local Similarity 37.8%; Pred. No. 62;

Matches 14; Conservative 7; Mismatches 9; Indels 7; Gaps 2;

QY 1 QLHNEGFKLFATEATSDWLNANNVPATPVAMPVPSQEQ 37

Db 180 ELHKGFKLWFDLIGLWVSQNLFP---QW---EQG 209

RESULT 14

ID Q9I6H0 PRELIMINARY; PRT; 344 AA.
AC Q9I6H0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable acetylpolymaine aminohydrolase.
GN PA0321.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004470; AAG03710.1; -.
DR InterPro; IPR000286; His_deacetylse.
DR Pfam; PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 344 AA; 37883 MW; 8238194D8911D0CE CRC64;

Query Match 25.6%; Score 69.5; DB 16; Length 344;
Best Local Similarity 37.1%; Pred. No. 1.6;
Matches 13; Conservative 7; Mismatches 14; Indels 1; Gaps 1;

QY 1 QLNHEGFKLFATEATSDWL-NANNVPATPVWFSQ 34
:|:|:| | | | | | : |:|:|:|:
Db 61 RVHSEGFVRFQNAWQDWLATGRSHDMLPIAWPTR 95

RESULT 9
O93937 PRELIMINARY; PRT; 2275 AA.
ID O93937
AC O93937;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE PyABCN (EC 6.3.5.5).
GN PyABCN.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=162425;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FGSC 4;
RA Aleksenko A., Liu W., Gojkovic Z., Nielsen J., Piskur J.;
RT "Structural and transcriptional analysis of the pyrABCN, pyrD and pyrF
RT genes in Aspergillus nidulans and the evolutionary origin of fungal
RT dihydroorotases.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AF112473; AAD09129.1; -.
DR HSSP; P00968; 1CS0.
DR InterPro; IPR006220; Anth_synthII.
DR InterPro; IPR006130; Asp/Orn_Cotranf.
DR InterPro; IPR002082; Asp_carbmltransf.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; Cpase_L.
DR InterPro; IPR005479; Cpase_L_D2.
DR InterPro; IPR005480; Cpase_L_D3.
DR InterPro; IPR005481; Cpase_L_N.
DR InterPro; IPR002474; CPSase_sm_chain.

DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR000991; GATase_1.
DR InterPro; IPR004362; MGS_like.
DR InterPro; IPR006131; OTCace_O.
DR InterPro; IPR006132; OTCace_P.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 3.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR Pfam; PF02142; MGS; 1.
DR Pfam; PF00185; OTCace; 1.
DR Pfam; PF02729; OTCace_N; 1.
DR PRINTS; PR00097; ANTSNTHASEII.
DR PRINTS; PR00100; AOTCASE.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR PRINTS; PR00096; GATASE.
DR TIGRFAMS; TIGR00670; asp_carb_tr; 1.
DR TIGRFAMS; TIGR01368; CPSaseIIsmall; 1.
DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
DR PROSITE; PS00442; GATASE_TYPE_I; 1.
KW Ligase.
SQ SEQUENCE 2275 AA; 249739 MW; 27BA9C1FA751436A CRC64;

Query Match 25.5%; Score 69; DB 3; Length 2275;
Best Local Similarity 52.0%; Pred. No. 16;
Matches 13; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 QLNHEGFKLFATEATSDWLNANNVP 25
:|:|:|:| | | | | | : |:|:|:|:
Db 1422 KLRDVGFELEATSGTADFLKENGVP 1446

RESULT 10
Q99ZV5 PRELIMINARY; PRT; 339 AA.
ID Q99ZV5
AC Q99ZV5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SPY1056.
GN SPY1056.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SE370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL; AE006550; AAK33943.1; -.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR004630; Cons_hypoth698.
DR Pfam; PF03601; Cons_hypoth698; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 339 AA; 36239 MW; A03E1866CF0DD99A CRC64;

Query Match 23.6%; Score 64; DB 16; Length 339;
Best Local Similarity 38.8%; Pred. No. 8;
Matches 19; Conservative 7; Mismatches 19; Indels 4; Gaps 2;

QY 2 LENEKGKLFATEATSDWLNANNVPATPVWFSQEQNP-SLSSIRKLIR 49

QY 1 QLNHEGFKLFATEATSDWLNANNVPATPVAVPWSQEQGNPSLSSIRKLRDG 51
|||:||||||| ||||| |||||:|:| |||||:|:| |||||:|:| |||||
Db 1381 QLHEQGKLYATEGTSWLNVDVPTAPVSWPTABDHSSSAPSTKLIHDG 1431

RESULT 6
Q9PTM9 PRELIMINARY; PRT; 1504 AA.
AC Q9PTM9;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Acetylglutamate-activated carbamoyl phosphate synthase III.
OS Opsanus beta.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Batrachoididae; Opsanus.
OX NCBI_TaxID=95145;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Kong H., Kahatapitiya N., Kingsley K., Salo W.L., Anderson P.M.,
RA Wang Y.S., Walsh P.J.;
RT "Induction of expression of carbamoyl phosphate synthetase III and
RT glutamine synthetase mRNA during confinement stress in gulf toadfish
RT (Opsanus beta).";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF169248; AAD51318.2; -
DR HSSP; P00968; 1A9X.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; CPase_L.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005480; CPase_L_D3.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR002474; CPSase_sm_chain.
DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR000991; GATase_1.
DR InterPro; IPR004362; MGS_like.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR TIGRFAMS; TIGR01368; CPSaseII_small; 1.
DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
SQ SEQUENCE 1504 AA; 165446 MW; C46D618F65AA3F8D CRC64;

Query Match 59.6%; Score 161.5; DB 13; Length 1504;
Best Local Similarity 62.7%; Pred. No. 2.1e-11;
Matches 32; Conservative 10; Mismatches 8; Indels 1; Gaps 1;

QY 1 QLNHEGFKLFATEATSDWLNANNVPATPVAVPWSQEQGNPSLSSIRKLRDG 51
|||:||||||| ||||| |||||:|:| |||||:|:| |||||:|:| |||||
Db 1380 QLKDEGFKLYATEATSAWLCANDVPSTPVAVPS-DNADSNLPSIKRLISEG 1429

RESULT 7
Q18990 PRELIMINARY; PRT; 2198 AA.
AC Q18990;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-JAN-1998 (TremBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE D2085.1 protein.
GN D2085.1.

OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Baynes C.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z54284; CAA91059.1; -
DR HSSP; P00479; 3CSU.
DR WormPep; D2085.1; CE03105.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR006220; Anth_synthII.
DR InterPro; IPR006130; Asp/Orn_Cotranf.
DR InterPro; IPR002082; Asp_carbmitransf.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; CPase_L.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005480; CPase_L_D3.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR002474; CPSase_sm_chain.
DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR002195; Dihydroorotase.
DR InterPro; IPR000991; GATase_1.
DR InterPro; IPR004362; MGS_like.
DR InterPro; IPR006131; OTCace_O.
DR InterPro; IPR006132; OTCace_P.
DR InterPro; IPR005847; Urease.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR Pfam; PF02142; MGS; 1.
DR Pfam; PF00185; OTCace; 1.
DR Pfam; PF02729; OTCace_N; 1.
DR PRINTS; PR00097; ANTSNTHASEII.
DR PRINTS; PR00100; AOTCASE.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR PRINTS; PR00096; GATASE.
DR PRODOM; PD000518; Urease; 1.
DR TIGRFAMS; TIGR00670; asp_carb_tr; 1.
DR TIGRFAMS; TIGR01368; CPSaseII_small; 1.
DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
DR PROSITE; PS00482; DIHYDROOROTASE_1; 1.
DR PROSITE; PS00483; DIHYDROOROTASE_2; 1.
DR PROSITE; PS00442; GATASE_TYPE_I; 1.
SQ SEQUENCE 2198 AA; 242567 MW; 2C029AFD34C71A4E CRC64;

Query Match 29.2%; Score 79; DB 5; Length 2198;
Best Local Similarity 33.3%; Pred. No. 0.86;
Matches 11; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 6 GFKLFATEATSDWLNANNVPATPVAVPWSQEQGN 38
|||:||||||| ||||| |||||:|:| |||||:|:| |||||
Db 1349 GYELGSKGTADYFQSNKINKVPVDWPFEEGSS 1381

RESULT 8
Q9I650

DR InterPro; IPR000991; GATase_1.
DR InterPro; IPR004362; MGS_like.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR TIGRFAMS; TIGR01368; CPSaseIISmall; 1.
DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
SQ SEQUENCE 1505 AA; 164592 MW; 89C4A22285166CA2 CRC64;

Query Match 61.3%; Score 166; DB 13; Length 1505;
Best Local Similarity 64.7%; Pred. No. 5.5e-12;
Matches 33; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 QLNHEGFKLFATEATSDWLNANNVPAIPVAPWPSQEGONPSLSSIRKLI RDG 51
|| |||||:||||| || |||||:||||| : || |||||: ||
Db 1380 QLKEEGFKLYATEATSAWLCANDVPAIPVAPWPTGMGDNLSPLSIKRLISEG 1430

RESULT 5
Q92115
ID Q92115 PRELIMINARY; PRT; 1502 AA.
AC Q92115; Q91438;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
DE Carbamoyl phosphate synthetase III [glutamine] mitochondrial precursor
DE (EC 6.3.5.5) (Glutamine-dependent carbamoyl-phosphate synthetase III)
DE (GD-CPSase III).
GN CPSIII.
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 39-47.
RC TISSUE=LIVER;
RX MEDLINE=95018247; PubMed=7932737;
RA Hong J., Salo W.L., Lusty C.J., Anderson P.M.;
RT "Carbamyl phosphate synthetase III, an evolutionary intermediate in
RT the transition between glutamine-dependent and ammonia-dependent
RT carbamyl phosphate synthetases.";
RL J. Mol. Biol. 243:131-140(1994).
RN [2]
RP SEQUENCE OF 1-42 FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=97149029; PubMed=8995057;
RA Hong J., Salo W.L., Chen Y., Atkinson B.G., Anderson P.M.;
RT "The promoter region of the carbamoyl-phosphate synthetase III gene of
RT Squalus acanthias.";
RL J. Mol. Evol. 43:602-609(1996).
RN [3]
RP ENZYME ACTIVITY, AND REGULATION.
RC TISSUE=LIVER;
RX MEDLINE=80168997; PubMed=6245445;
RA Anderson P.M.;
RT "Glutamine- and N-acetylglutamate-dependent carbamoyl phosphate
RT synthetase in elasmobranchs.";
RL Science 208:291-293(1980).
RN [4]
RP ENZYME ACTIVITY, AND SUBUNITS.
RC TISSUE=LIVER;
RX MEDLINE=82053090; PubMed=7298655;
RA Anderson P.M.;
RT "Purification and properties of the glutamine- and N-acetyl-L-
RT glutamate-dependent carbamoyl phosphate synthetase from liver of
RT Squalus acanthias.";

RL InterPro; IPR000991; GATase_1.
RN [5]
RP SUBCELLULAR LOCATION.
RC TISSUE=LIVER;
RX MEDLINE=82214067; PubMed=6123510;
RA Casey C.A., Anderson P.M.;
RT "Subcellular location of glutamine synthetase and urea cycle enzymes
RT in liver of spiny dogfish (Squalus acanthias).";
RL J. Biol. Chem. 257:8449-8453(1982).
CC -!- FUNCTION: INVOLVED IN THE UREA CYCLE OF UREOTELIC ANIMALS WHERE
CC THE ENZYME PLAYS AN IMPORTANT ROLE IN THE GLUTAMINE-DEPENDENT
CC FORMATION OF CARBAMOYL PHOSPHATE FOR UREA SYNTHESIS. UREA MAYBE
CC RETAINED IN TISSUES AND BODY FLUIDS AS A MECHANISM OF
CC OSMOREGULATION.
CC -!- CATALYTIC ACTIVITY: 2 ATP + GLUTAMINE + CO(2) + H(2)O = 2 ADP +
CC PHOSPHATE + GLUTAMATE + CARBAMOYL PHOSPHATE.
CC -!- ENZYME REGULATION: REQUIRES N-ACETYLGLUTAMATE AS AN ALLOSTERIC
CC ACTIVATOR.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LIVER. MAY NOT BE
CC EXPRESSED IN OTHER TISSUES.
CC -!- SIMILARITY: TO OTHER CARBAMOYL-PHOSPHATE SYNTHETASES. ALSO
CC CONTAINS A GLUTAMINE AMIDOTRANSFERASE-LIKE DOMAIN. THE CYSTEINE
CC RESIDUE WHICH IS ESSENTIAL FOR AMIDO TRANSFERASE ACTIVITY, IS
CC CONSERVED.
DR EMBL; L31362; AAA96435.1; -.
DR EMBL; U19771; AAB49032.1; -.
DR HSSP; P00968; ICS0.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; CPase_L.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005480; CPase_L_D3.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR002474; CPSase_sm_chain.
DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR000991; GATase_1.
DR InterPro; IPR004362; MGS_like.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR PRINTS; PR00096; GATASE.
DR TIGRFAMS; TIGR01368; CPSaseIISmall; 1.
DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
KW Ligase; Transit peptide; Mitochondrion; ATP-binding; Urea cycle.
FT TRANSIT 1 38 MITOCHONDRION.
FT CHAIN 39 1502 CARBAMOYL PHOSPHATE SYNTHETASE III
FT [GLUTAMINE] MITOCHONDRIAL.
FT ANTHRANILATE PHOSPHORIBOSYLTRANSFERASE-
FT LIKE.
FT DOMAIN 39 219 GLUTAMINE AMIDOTRANSFERASE-LIKE.
FT DOMAIN 220 411 CARBAMOYL-PHOSPHATE SYNTHETASE LARGE
FT DOMAIN 412 1502 CHAIN.
FT NP_BIND 572 627 ATP (2) (POTENTIAL).
FT NP_BIND 719 769 ATP (1) (POTENTIAL).
FT NP_BIND 1114 1172 ATP (2) (POTENTIAL).
FT NP_BIND 1259 1303 ATP (1) (POTENTIAL).
FT REPEAT 420 877
FT REPEAT 971 1411
FT ACT_SITE 294 294 GATASE (BY SIMILARITY).
SQ SEQUENCE 1502 AA; 165140 MW; CF444C222D99E2EE CRC64;

Query Match 60.9%; Score 165; DB 13; Length 1502;
Best Local Similarity 56.9%; Pred. No. 7.4e-12;
Matches 29; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Carbamoyl-phosphate synthetase III (EC 6.3.5.5).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97197794; PubMed=9045644;
RA Korte J.J., Salo W.L., Cabrera V.M., Wright P.A., Felskie A.K.,
RA Anderson P.M.;
RT "Expression of carbamoyl-phosphate synthetase III mRNA during the
early stages of development and in muscle of adult rainbow trout
(Oncorhynchus mykiss).";
RL J. Biol. Chem. 272:6270-6277(1997).
DR EMBL; U65893; AAC60207.1; -.
DR HSSP; P00968; 1CS0.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; CPase_L.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005480; CPase_L_D3.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR002474; CPase_sm_chain.
DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR000991; GATase_1.
DR InterPro; IPR004362; MGS_like.
DR Pfam; PF0289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR PRINTS; PR00096; GATASE.
DR TIGRFAMS; TIGR01368; CPSaseIIsml; 1.
DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
KW Ligase.
SQ SEQUENCE 1518 AA; 166577 MW; F9D6331C710A0EBD CRC64;

Query Match 72.0%; Score 195; DB 13; Length 1518;
Best Local Similarity 74.5%; Pred. No. 1.2e-15;
Matches 38; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 QLNHEGKLFATEATSDWLNANNVPATPVAVPWSQEGQNPSSLIRKLRDG 51
||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : || || || : ||
Db 1378 QLNHEGKLFATEATSAWLSANDVPATPVAVPWSQEGGDASLPSIIRLINEG 1428

RESULT 3
O42433
ID O42433 PRELIMINARY; PRT; 1506 AA.
AC O42433;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Carbamoyl-phosphate synthetase III (EC 6.3.5.5).
OS Micropterus salmoides (Largemouth bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Centrarchidae; Micropterus.
OX NCBI_TaxID=27706;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Kong H., Edberg D.D., Salo W.L., Korte J.J., Wright P.A.,
RA Anderson P.M.;

RT "Nitrogen Excretion and Expression of Carbamoyl-Phosphate Synthetase
III in Extra-Hepatic Tissues of Largemouth Bass (Micropterus
salmoides).";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF006491; AAB62566.1; -.
DR HSSP; P00968; 1CS0.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; CPase_L.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005480; CPase_L_D3.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR002474; CPase_sm_chain.
DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR000991; GATase_1.
DR InterPro; IPR004362; MGS_like.
DR Pfam; PF0289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR PRINTS; PR00096; GATASE.
DR TIGRFAMS; TIGR01368; CPSaseIIsml; 1.
DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
DR PROSITE; PS00866; CPSASE_1; 1.
DR PROSITE; PS00867; CPSASE_2; 2.
KW Ligase.
SQ SEQUENCE 1505 AA; 165304 MW; 59439669AC8BB86D CRC64;

Query Match 64.9%; Score 176; DB 13; Length 1506;
Best Local Similarity 66.0%; Pred. No. 3e-13;
Matches 33; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 2 LHNKGKLFATEATSDWLNANNVPATPVAVPWSQEGQNPSSLIRKLRDG 51
| ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| : || |||| : ||
Db 1382 LKEGFKLYATEATSAWLCANDVPATPVAVPTEKGGDTSLPSIKRLINEG 1431

RESULT 4
Q9W7F3
ID Q9W7F3 PRELIMINARY; PRT; 1505 AA.
AC Q9W7F3;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Carbamoyl-phosphate synthetase III.
OS Alcolapia grahami.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Alcolapia.
OX NCBI_TaxID=87886;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99445540; PubMed=10514466;
RA Lindley T.E., Scheiderer C.L., Walsh P.J., Wood C.M., Bergman H.L.,
RA Bergman A.L., Laurent P., Wilson P., Anderson P.M.;
RT "Muscle as the primary site of urea cycle enzyme activity in an
alkaline lake-adapted tilapia, Oreochromis alcalicus grahami.";
RL J. Biol. Chem. 274:29858-29861(1999).
DR EMBL; AF119250; AAD43968.1; -.
DR HSSP; P00968; 1A9X.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; CPase_L.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005480; CPase_L_D3.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR002474; CPSase_sm_chain.
DR InterPro; IPR001317; CPS_GATase.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: September 13, 2003, 02:07:12 ; Search time 57.5 Seconds
(without alignments)
228.881 Million cell updates/sec

Title: US-09-585-077C-4_COPY_1380_1430
Perfect score: 271
Sequence: 1 QLNHEGFKLFATEATSDWLN.....PSQEGQNPSLSIRKLIRDG 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_23:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertibrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archheap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	271	100.0	748	11 Q8C196	Q8c196 mus musculu
2	195	72.0	1518	13 P70087	P70087 oncorhynchu
3	176	64.9	1506	13 O42433	O42433 micropterus
4	166	61.3	1505	13 Q9W7F3	Q9w7f3 alcolapia g
5	165	60.9	1502	13 Q92115	Q92115 squalus aca
6	161.5	59.6	1504	13 Q9PTM9	Q9ptm9 opsanus bet
7	79	29.2	2198	5 Q18990	Q18990 caenorhabdi
8	69.5	25.6	344	16 Q9I6H0	Q9i6h0 pseudomonas
9	69	25.5	2275	3 O93937	O93937 emericella
10	64	23.6	339	16 Q99ZV5	Q99zv5 streptococc
11	64	23.6	339	16 Q8K7M5	Q8k7m5 streptococc
12	64	23.6	863	16 Q8UG58	Q8ug58 agrobacteri
13	63.5	23.4	1798	4 Q14997	Q14997 homo sapien
14	63.5	23.4	1803	11 Q9ESY1	Q9esy1 rattus norv
15	63.5	23.4	2921	5 Q9N973	Q9n973 leishmania
16	63	23.2	339	16 Q8P171	Q8p171 streptococc

17	63	23.2	540	3 Q00878	Q00878 nectria hae
18	63	23.2	1065	16 Q8EZX0	Q8ezx0 leptospira
19	63	23.2	2204	12 Q99FK6	Q99fk6 porcine tes
20	62	22.9	336	16 Q97TA9	Q97ta9 streptococc
21	62	22.9	340	16 Q8DRN6	Q8drn6 streptococc
22	62	22.9	3564	11 Q923L3	Q923l3 mus musculu
23	61.5	22.7	1068	16 Q8CXH7	Q8cxh7 oceanobacil
24	61	22.5	185	5 Q8IPB6	Q8ipb6 drosophila
25	61	22.5	510	5 Q9VKT5	Q9vkt5 drosophila
26	60.5	22.3	246	4 Q9UJ18	Q9uj18 homo sapien
27	60.5	22.3	1360	12 Q55253	Q55253 murine hepa
28	60	22.1	157	16 Q8P5C0	Q8p5c0 xanthomonas
29	60	22.1	341	16 Q8DRZ4	Q8drz4 streptococc
30	60	22.1	455	16 Q8PJ84	Q8pj84 xanthomonas
31	60	22.1	889	2 Q8KNY4	Q8kny4 xanthomonas
32	60	22.1	1113	16 Q8FT42	Q8ft42 corynebacte
33	59.5	22.0	508	4 Q8IZH4	Q8izh4 homo sapien
34	59.5	22.0	537	16 Q9A8M4	Q9a8m4 caulobacter
35	59.5	22.0	1181	4 Q9UM06	Q9um06 homo sapien
36	59.5	22.0	1220	4 Q9UM05	Q9um05 homo sapien
37	59.5	22.0	1220	4 Q94979	Q94979 homo sapien
38	59.5	22.0	1249	11 Q922Q1	Q922q1 rattus norv
39	59.5	22.0	1698	12 Q9YLR3	Q9ylr3 hepatitis e
40	59	21.8	375	10 Q8VXX9	Q8vxx9 fagus sylv
41	59	21.8	523	16 Q8PKJ9	Q8pkj9 xanthomonas
42	59	21.8	1012	11 Q9WVM6	Q9wvm6 mus musculu
43	58.5	21.6	277	16 Q92DF9	Q92df9 listeria in
44	58.5	21.6	277	16 Q8Y8N5	Q8y8n5 listeria mo
45	58.5	21.6	715	3 Q12204	Q12204 saccharomyc

ALIGNMENTS

RESULT 1

Q8C196
ID Q8C196 PRELIMINARY; PRT; 748 AA.
AC Q8C196;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Weakly similar to carbamoyl-phosphate synthase (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK028683; BAC26064.1; -.
FT NON_TER 1
SQ SEQUENCE 748 AA; 83277 MW; AA818F3BAB82D9A1 CRC64;

Query Match 100.0%; Score 271; DB 11; Length 748;
Best Local Similarity 100.0%; Pred. No. 1.3e-25;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLNHEGFKLFATEATSDWLNANNVPATVPVWPSQEGQNPSLSIRKLIRDG 51
|||
Db 628 QLNHEGFKLFATEATSDWLNANNVPATVPVWPSQEGQNPSLSIRKLIRDG 678

RESULT 2

P70087
ID P70087 PRELIMINARY; PRT; 1518 AA.
AC P70087;
DT 01-FEB-1997 (TREMELrel. 02, Created)

Db 1335 LESGLSYLASLGTADFYTEHGVKVTAVDWHFEEAVDGECPQRSI 1380

RESULT 12

PYRI_MESAU

ID PYRI_MESAU STANDARD; PRT; 2225 AA.

AC P08955; P70108;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE CAD protein [Includes: Glutamine-dependent carbamoyl-phosphate synthase (EC 6.3.5.5); Aspartate carbamoyltransferase (EC 2.1.1.3.2); Dihydroorotase (EC 3.5.2.3)].

DE GN CAD.

OS Mesocricetus auratus (Golden hamster).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Mesocricetus.

OX NCBI_TaxID=10036;

RN [1]

RP SEQUENCE OF 1-169 FROM N.A.

RX MEDLINE=91139675; PubMed=1671675;

RA Bein K., Simmer J.P., Evans D.R.;

RT "Molecular cloning of a cDNA encoding the amino end of the mammalian multifunctional protein CAD and analysis of the 5'-flanking region of the CAD gene.";

RL J. Biol. Chem. 266:3791-3799(1991).

RN [2]

RP SEQUENCE OF 156-1455 FROM N.A.

RX MEDLINE=90285162; PubMed=1972379;

RA Simmer J.P., Kelly R.E., Rinker A.G. Jr., Scully J.L., Evans D.R.;

RT "Mammalian carbamyl phosphate synthetase (CPS). DNA sequence and evolution of the CPS domain of the Syrian hamster multifunctional protein CAD.";

RL J. Biol. Chem. 265:10395-10402(1990).

RN [3]

RP SEQUENCE OF 246-513 FROM N.A.

RX MEDLINE=88309082; PubMed=2900634;

RA Maley J.A., Davidson J.N.;

RT "Identification of the junction between the glutamine amidotransferase and carbamyl phosphate synthetase domains of the mammalian CAD protein.";

RL Biochem. Biophys. Res. Commun. 154:1047-1053(1988).

RN [4]

RP SEQUENCE OF 1391-1870 FROM N.A.

RX MEDLINE=90115834; PubMed=1967494;

RA Simmer J.P., Kelly R.E., Rinker A.G. Jr., Zimmermann B.H., Scully J.L., Kim H., Evans D.R.;

RT "Mammalian dihydroorotase: nucleotide sequence, peptide sequences, and evolution of the dihydroorotase domain of the multifunctional protein CAD.";

RL Proc. Natl. Acad. Sci. U.S.A. 87:174-178(1990).

RN [5]

RP SEQUENCE OF 1774-2225 FROM N.A.

RX MEDLINE=89282776; PubMed=2543974;

RA Simmer J.P., Kelly R.E., Scully J.L., Grayson D.R., Rinker A.G. Jr., Bergh S.T., Evans D.R.;

RT "Mammalian aspartate transcarbamylase (ATCase): sequence of the ATCase domain and interdomain linker in the CAD multifunctional polypeptide and properties of the isolated domain.";

RL Proc. Natl. Acad. Sci. U.S.A. 86:4382-4386(1989).

RN [6]

RP SEQUENCE OF 2074-2225 FROM N.A.

RX MEDLINE=85267690; PubMed=2862577;

RA Shigesada K., Stark G.R., Maley J.A., Niswander L.A., Davidson J.N.;

RT "Construction of a cDNA to the hamster CAD gene and its application toward defining the domain for aspartate transcarbamylase.";

RL Mol. Cell. Biol. 5:1735-1742(1985).

RN [7]

RP DOMAINS, AND PARTIAL SEQUENCE.

RX MEDLINE=92202288; PubMed=1348059;

RA Kim H., Kelly R.E., Evans D.R.;

RT "The structural organization of the hamster multifunctional protein

RT CAD. Controlled proteolysis, domains, and linkers.";

RL J. Biol. Chem. 267:7177-7184(1992).

RN [8]

RX 3D-STRUCTURE MODELING OF ATCASE DOMAIN.

RA MEDLINE=91172744; PubMed=2006137;

RA Scully J.L., Evans D.R.;

RT "Comparative modeling of mammalian aspartate transcarbamylase.";

RL Proteins 9:191-206(1991).

CC -!- FUNCTION: THIS PROTEIN IS A "FUSION" PROTEIN ENCODING FOUR ENZYMATIC ACTIVITIES OF THE PYRIMIDINE PATHWAY (GATASE, CPSASE, ATCASE AND DHOASE).

CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP + phosphate + L-glutamate + carbamoyl phosphate.

CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate = phosphate + N-carbamoyl-L-aspartate.

CC -!- CATALYTIC ACTIVITY: (S)-dihydroorotate + H(2)O = N-carbamoyl-L-aspartate.

CC -!- Cofactor: DHOASE ACTIVITY REQUIRES A ZINC ATOM.

CC -!- ENZYME REGULATION: ALLOSTERICALLY REGULATED AND CONTROLLED BY PHOSPHORYLATION. PRPP IS AN ACTIVATOR WHILE UMP IS AN INHIBITOR OF THE CPSASE REACTION.

CC -!- PATHWAY: FIRST TO THIRD STEP IN PYRIMIDINE BIOSYNTHESIS.

CC -!- SUBUNIT: Homohexamer.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- MISCELLANEOUS: GATASE (GLUTAMINE AMIDOTRANSFERASE) AND CPSASE (CARBAMOYL PHOSPHATE SYNTHASE) FORM TOGETHER THE GLUTAMINE-DEPENDENT CPSASE (GD-CPSASE) (EC 6.3.5.5).

CC -!- SIMILARITY: THE CPSASE DOMAIN IS SIMILAR TO OTHER CPASES.

CC -!- SIMILARITY: IN THE CENTRAL SECTION; BELONGS TO THE DHOASE FAMILY.

CC -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

CC -----

DR EMBL; J05503; AAA37062.1; -.

DR EMBL; M28866; AAA37073.1; -.

DR EMBL; M60078; AAA63617.1; -.

DR EMBL; M11242; AAA37061.1; -.

DR EMBL; M23652; AAA37064.1; -.

DR EMBL; M21927; AAA37063.1; -.

DR PIR; A38653; A23443.

DR HSSP; P00479; 3CSU.

DR InterPro; IPR006680; Amidohydro_1.

DR InterPro; IPR006220; Anth_synthII.

DR InterPro; IPR006130; Asp/Orn_Cotranf.

DR InterPro; IPR002082; Asp_carbmltransf.

DR InterPro; IPR006275; CarA_L_glu.

DR InterPro; IPR006274; CarA_small.

DR InterPro; IPR005483; CPase_L.

DR InterPro; IPR005479; CPase_L_D2.

DR InterPro; IPR005480; CPase_L_D3.

DR InterPro; IPR005481; CPase_L_N.

DR InterPro; IPR001317; CPS_GATase.

DR InterPro; IPR002474; CPSase_sm_chain.

DR InterPro; IPR002195; Dihydroorotase.

DR InterPro; IPR000991; GATase_1.

DR InterPro; IPR004362; MGS_like.

DR InterPro; IPR006131; OTCace_O.

DR InterPro; IPR006132; OTCace_P.

DR InterPro; IPR004722; Pyrc_multi.

DR InterPro; IPR005847; Urease.

DR Pfam; PF01979; Amidohydro_1; 1.

DR Pfam; PF00289; CPSase_L_chain; 2.

DR Pfam; PF02786; CPSase_L_D2; 2.

DR Pfam; PF02787; CPSase_L_D3; 1.

DR Pfam; PF00988; CPSase_sm_chain; 1.

DR Pfam; PF00117; GATase; 1.

DR Pfam; PF02142; MGS; 1.

RESULT 10
PYR1_SCHPO STANDARD; PRT; 2244 AA.
AC Q09794;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE URAl protein [includes: Glutamine-dependent carbamoyl-phosphate
synthase (EC 6.3.5.5); Aspartate carbamoyltransferase (EC 2.1.3.2)].
GN URAl OR SPAC2267.06C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE OF 22-2244 FROM N.A.
RC STRAIN=972;
RX MEDLINE=96020160; PubMed=8590465;
RA Lollier M., Jaquet L., Nedeva T., Lacroute F., Potier S.,
RA Souciet J.-L.;
RT "As in Saccharomyces cerevisiae, aspartate transcarbamoylase is
RT assembled on a multifunctional protein including a dihydroorotase-like
RT cryptic domain in Schizosaccharomyces pombe.";
RL Curr. Genet. 28:138-149(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=2184801; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham S., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadiou E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: THIS PROTEIN IS A "FUSION" PROTEIN ENCODING THREE
CC ENZYMATIC ACTIVITIES OF THE PYRIMIDINE PATHWAY (GATASE, CPSASE,
CC AND ATCase) (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + L-glutamate + carbamoyl phosphate.
CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate = phosphate
CC + N-carbamoyl-L-aspartate.
CC -!- PATHWAY: Pyrimidine biosynthesis; first step.
CC -!- PATHWAY: Pyrimidine biosynthesis; second step.
CC -!- MISCELLANEOUS: GATASE (GLUTAMINE AMIDOTRANSFERASE) AND CPSASE
CC (CARBAMOYL PHOSPHATE SYNTHASE) FORM TOGETHER THE
CC GLUTAMINE-DEPENDENT CPSASE (GD-CPSASE) (EC 6.3.5.5).
CC -!- MISCELLANEOUS: IN EUKARYOTES EC 6.3.5.5 IS SYNTHESIZED BY TWO
CC PATHWAY-SPECIFIC (ARGININE AND PYRIMIDINE) UNDER SEPARATE CONTROL.
CC -!- SIMILARITY: THE CPSASE DOMAIN IS SIMILAR TO OTHER CPASES.
CC -!- SIMILARITY: IN THE CENTRAL SECTION; BELONGS TO THE DHOASE FAMILY.
CC -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.
CC

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X81841; CAA57433.1; -.
EMBL; Z54328; CAA91130.1; -.
PIR; S65074; S65074.
PIR; T11616; T11616.
HSSP; P00968; 1A9X.
GeneDB_SPombe; SPAC2267.06c; -.
InterPro; IPR006220; Anth_synthII.
InterPro; IPR006130; Asp/Orn_Cotranf.
InterPro; IPR002082; Asp_carbamtransf.
InterPro; IPR006275; CarA_L_glu.
InterPro; IPR006274; CarA_small.
InterPro; IPR005483; Cpase_L.
InterPro; IPR005479; Cpase_L_D2.
InterPro; IPR005480; Cpase_L_D3.
InterPro; IPR005481; Cpase_L_N.
InterPro; IPR001317; CPS_GATase.
InterPro; IPR002474; CPSase_sm_chain.
InterPro; IPR000991; GATase_1.
InterPro; IPR004362; MGS_like.
InterPro; IPR006131; OTCace_O.
InterPro; IPR006132; OTCace_P.
Pfam; PF00289; CPSase_L_chain; 2.
Pfam; PF02786; CPSase_L_D2; 2.
Pfam; PF02787; CPSase_L_D3; 1.
Pfam; PF00988; CPSase_sm_chain; 1.
Pfam; PF00117; GATase; 1.
Pfam; PF02142; MGS; 1.
Pfam; PF00185; OTCace; 1.
Pfam; PF02729; OTCace_N; 1.
PRINTS; PR00097; ANTSNTHASEII.
PRINTS; PR00100; AOTCase.
PRINTS; PR00098; CPSASE.
PRINTS; PR00099; CPSGATASE.
PRINTS; PR00096; GATASE.
TIGRFAMS; TIGR00670; asp_carb_tr; 1.
TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
TIGRFAMS; TIGR01368; CPSaseII_small; 1.
PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
PROSITE; PS00442; GATASE_TYPE_I; 1.
PROSITE; PS00866; CPSASE_1; 2.
PROSITE; PS00867; CPSASE_2; 2.
KW Pyrimidine biosynthesis; Ligase; Transferase; Multifunctional enzyme.
FT DOMAIN 1 437 GATASE (GLUTAMINE AMIDOTRANSFERASE)
FT (BY SIMILARITY).
FT DOMAIN 438 477 LINKER (BY SIMILARITY).
FT DOMAIN 478 1514 CPSASE (CARBAMOYL-PHOSPHATE SYNTHASE)
FT (BY SIMILARITY).
FT DOMAIN 1515 1524 LINKER (BY SIMILARITY).
FT DOMAIN 1525 1853 DEFECTIVE DHOASE DOMAIN
FT (BY SIMILARITY).
FT DOMAIN 1854 1935 LINKER (BY SIMILARITY).
FT DOMAIN 1936 2244 ATCase (ASPARTATE TRANSCARBAMYLASE)
FT (BY SIMILARITY).
FT ACT_SITE 338 338 GATASE (BY SIMILARITY).
FT ACT_SITE 422 422 GATASE (BY SIMILARITY).
FT ACT_SITE 424 424 GATASE (BY SIMILARITY).
FT CONFLICT 336 338 GIC -> RYF (IN REF. 1).
FT CONFLICT 1035 1039 CAVRA -> LQFAQ (IN REF. 1).
FT CONFLICT 1409 1410 EL -> DV (IN REF. 1).
FT CONFLICT 1975 1975 G -> E (IN REF. 1).
FT CONFLICT 2002 2002 G -> E (IN REF. 1).
SQ SEQUENCE 2244 AA; 248306 MW; 5700D153B50CD3E9 CRC64;

Query Match 23.2%; Score 63; DB 1; Length 2244;

ID	CARB_BACSU	STANDARD;	PRT;	1071 AA.
AC	P25994;			
DT	01-MAY-1992 (Rel. 22, Created)			
DT	01-MAY-1992 (Rel. 22, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Carbamoyl-phosphate synthase, pyrimidine-specific, large chain			
DE	(EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).			
GN	PYRAB.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1423;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91225016; PubMed=1709162;			
RA	Quinn C.L., Stephenson B.T., Switzer R.L.;			
RA	"Functional organization and nucleotide sequence of the Bacillus			
RT	subtilis pyrimidine biosynthetic operon.";			
RT	J. Biol. Chem. 266:9113-9127(1991).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=168;			
RX	MEDLINE=98044033; PubMed=9384377;			
RA	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,			
RA	Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,			
RA	Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,			
RA	Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,			
RA	Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,			
RA	Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,			
RA	Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,			
RA	Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,			
RA	Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,			
RA	Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,			
RA	Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,			
RA	Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,			
RA	Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,			
RA	Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,			
RA	Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,			
RA	Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,			
RA	Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,			
RA	Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,			
RA	Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,			
RA	Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,			
RA	Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,			
RA	Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,			
RA	Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,			
RA	Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,			
RA	Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,			
RA	Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,			
RA	Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,			
RA	Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;			
RT	"The complete genome sequence of the Gram-positive bacterium Bacillus			
RT	subtilis.";			
RL	Nature 390:249-256(1997).			
CC	-!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +			
CC	phosphate + L-glutamate + carbamoyl phosphate.			
CC	-!- COFACTOR: Binds 3 manganese ions per subunit (By similarity).			
CC	-!- PATHWAY: Pyrimidine biosynthesis; first step.			
CC	-!- SUBUNIT: Composed of two chains; the small (or glutamine) chain			
CC	promotes the hydrolysis of glutamine to ammonia, which is used by			
CC	the large (or ammonia) chain to synthesize carbamoyl phosphate.			
CC	-!- SIMILARITY: BELONGS TO THE CARB FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M59757; AAA21270.1; -.			
DR	EMBL; Z99112; CAB13426.1; -.			
DR	PIR; F39845; F39845.			

CC symptoms are vomiting in infancy, protein intolerance,
CC intermittent ataxia, seizures, lethargy, and mental retardation.
CC -!- SIMILARITY: TO OTHER CARBAMOYL-PHOSPHATE SYNTHETASES. ALSO
CC CONTAINS A GLUTAMINE AMIDOTRANSFERASE-LIKE DOMAIN BUT THE CYSTEINE
CC RESIDUE WHICH IS ESSENTIAL FOR AMIDO TRANSFERASE ACTIVITY, IS
CC MISSING IN THIS PROTEIN SO THIS DOMAIN IS NO LONGER ACTIVE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90282; BAA14328.1; -.
DR EMBL; Y15793; CAA75785.1; -.
DR EMBL; AF154830; AAD38072.1; -.
DR PIR; JQ1348; JQ1348.
DR HSSP; P00968; 1A9X.
DR Genew; HGNC:2323; CPS1.
DR GK; P31327; -.
DR MIM; 237300; -.
DR GO; GO:0005739; C:mitochondrion; TAS.
DR GO; GO:0004087; F:carbamoyl-phosphate synthase (ammonia) acti. . .; TAS.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; CPase_L.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005480; CPase_L_D3.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR002474; CPSase_sm_chain.
DR InterPro; IPR000991; GATase_1.
DR InterPro; IPR004362; MGS_like.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR PRINTS; PR00096; GATASE.
DR TIGRFAMs; TIGR01369; CPSaseII_lrg; 1.
DR TIGRFAMs; TIGR01368; CPSaseII_small; 1.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
KW Ligase; Allosteric enzyme; Repeat; Transit peptide; Mitochondrion;
KW ATP-binding; Urea cycle; Polymorphism; Disease mutation.
FT TRANSIT 1 38 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 39 1500 CARBAMOYL-PHOSPHATE SYNTHASE [AMMONIA].
FT NP_BIND 571 626 ATP (2).
FT NP_BIND 718 768 ATP (1).
FT NP_BIND 1113 1171 ATP (2).
FT NP_BIND 1259 1302 ATP (1).
FT DOMAIN 39 219 ANTHRANILATE PHOSPHORIBOSYLTRANSFERASE
FT HOMOLOG.
FT DOMAIN 220 410 GLUTAMINE AMIDOTRANSFERASE-LIKE.
FT DOMAIN 411 1500 CARBAMOYL-PHOSPHATE SYNTHETASE LARGE
FT CHAIN.
FT REPEAT 419 876
FT REPEAT 970 1410
FT VARIANT 337 337
FT VARIANT 344 344
FT VARIANT 344 344
FT VARIANT 544 544
FT VARIANT 544 544
FT CONFLICT 111 111
FT CONFLICT 279 279
FT CONFLICT 338 338
FT CONFLICT 718 722

FT CONFLICT 729 729 A -> T (IN REF. 1).
FT CONFLICT 749 749 E -> G (IN REF. 1).
FT CONFLICT 1161 1162 EH -> AT (IN REF. 1).
FT CONFLICT 1204 1205 GD -> EN (IN REF. 1).
FT CONFLICT 1254 1254 I -> N (IN REF. 1).
FT CONFLICT 1266 1266 F -> S (IN REF. 1).
FT CONFLICT 1283 1283 M -> L (IN REF. 1).
FT CONFLICT 1303 1303 A -> V (IN REF. 1).
FT CONFLICT 1406 1406 T -> N (IN REF. 1).
SQ SEQUENCE 1500 AA; 164939 MW; E53A22D77563961D CRC64;

Query Match 100.0%; Score 271; DB 1; Length 1500;
Best Local Similarity 100.0%; Pred. No. 2.8e-25;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLNHEGFKLFATEATSDWLNANNVPATPVAMPVPSQEQNPSSLSSIRKLI RDG 51
|||
Db 1380 QLNHEGFKLFATEATSDWLNANNVPATPVAMPVPSQEQNPSSLSSIRKLI RDG 1430

RESULT 2
CPSM_RAT
ID CPSM_RAT STANDARD; PRT; 1500 AA.
AC P07756;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carbamoyl-phosphate synthase [ammonia], mitochondrial precursor
DE (EC 6.3.4.16) (Carbamoyl-phosphate synthetase I) (CPSASE I).
GN CPS1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85261323; PubMed=2991241;
RA Nyunoya H., Broglie K.E., Widgren E.E., Lusty C.J.;
RT "Characterization and derivation of the gene coding for mitochondrial
RT carbamyl phosphate synthetase I of rat.";
RL J. Biol. Chem. 262:10415-10418(1987).
RN [2]
RP SEQUENCE OF 1-42 FROM N.A.
RX MEDLINE=87280088; PubMed=3038878;
RA Lagace M., Howell B.W., Burak R., Lusty C.J., Shore G.C.;
RT "Rat carbamyl-phosphate synthetase I gene. Promoter sequence and
RT tissue-specific transcriptional regulation in vitro.";
RL J. Biol. Chem. 262:10415-10418(1987).
CC -!- FUNCTION: INVOLVED IN THE UREA CYCLE OF UREOTELIC ANIMALS WHERE
CC THE ENZYME PLAYS AN IMPORTANT ROLE IN REMOVING EXCESS AMMONIA
CC FROM THE CELL.
CC -!- CATALYTIC ACTIVITY: 2 ATP + NH(3) + CO(2) + H(2)O = 2 ADP +
CC phosphate + carbamoyl phosphate.
CC -!- ENZYME REGULATION: REQUIRES N-ACETYLGLUTAMATE AS AN ALLOSTERIC
CC ACTIVATOR.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- TISSUE SPECIFICITY: PRIMARILY IN THE LIVER AND SMALL INTESTINE.
CC -!- PTM: 50% OF THE MATURE PROTEIN THAT WAS ISOLATED HAD LEU 39 AS ITS
CC N-TERMINAL RESIDUE AND 50% HAD SER 40 SUGGESTING TWO ADJACENT
CC PROCESSING SITES. HOWEVER, THE POSSIBILITY OF PROTEOLYTIC REMOVAL
CC OF LEU 39 DURING THE ISOLATION OF THE ENZYME CANNOT BE EXCLUDED.
CC -!- SIMILARITY: TO OTHER CARBAMOYL-PHOSPHATE SYNTHETASES. ALSO
CC CONTAINS A GLUTAMINE AMIDOTRANSFERASE-LIKE DOMAIN BUT THE CYSTEINE
CC RESIDUE WHICH IS ESSENTIAL FOR AMIDO TRANSFERASE ACTIVITY, IS
CC MISSING IN THIS PROTEIN SO THIS DOMAIN IS NO LONGER ACTIVE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2003, 02:03:24 ; Search time 14 Seconds
(without alignments)
171.312 Million cell updates/sec

Title: US-09-585-077C-4_COPY_1380_1430
Perfect score: 271
Sequence: 1 QLNHEGFKLFATEATSDWLN.....PSQEGQNPSLSIRKLIRDG 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	271	100.0	1500	1	CPSM_HUMAN
2	271	100.0	1500	1	CPSM_RAT
3	184.5	68.1	1496	1	CPSM_RANCA
4	82	30.3	618	1	CAB2_METJA
5	74.5	27.5	2185	1	PYR1_DICDI
6	69.5	25.6	1115	1	CARB_MYCTU
7	69	25.5	1102	1	CARB_STRCO
8	68.5	25.3	1071	1	CARB_BACSU
9	65	24.0	2214	1	PYR1_YEAST
10	63	23.2	2244	1	PYR1_SCHPO
11	59.5	22.0	2225	1	PYR1_HUMAN
12	59.5	22.0	2225	1	PYR1_MESAU
13	58	21.4	314	1	APAH_XANAC
14	58	21.4	452	1	PUR9_THEMA
15	58	21.4	591	1	ALU8_HUMAN
16	57.5	21.2	264	1	CBIK_SALTY
17	57.5	21.2	457	1	BAG4_MOUSE
18	57	21.0	2242	1	PYR1_SQUAC
19	56.5	20.8	811	1	REF1_YEAST
20	56.5	20.8	1121	1	CARB_MYCLE
21	56	20.7	1162	1	CARB_BRUME
22	55	20.3	209	1	LOLB_RALSO
23	55	20.3	254	1	VSPA_SOYBN
24	55	20.3	291	1	S25K_SOYBN
25	55	20.3	441	1	PAC1_MOUSE
26	55	20.3	441	1	PAC1_RAT
27	55	20.3	444	1	PAC1_HUMAN
28	55	20.3	521	1	YF92_MYCPN
29	55	20.3	1113	1	CARB_CORGL
30	54.5	20.1	449	1	VPI_CAV82
31	54.5	20.1	514	1	GSH1_BRAJU
32	54.5	20.1	662	1	YFL8_YEAST
33	54.5	20.1	1062	1	CARB_BACHD

34	54.5	20.1	1228	1	SLAP_BACST	P35825 bacillus st
35	54	19.9	423	1	EPD1_RALSO	P58591 ralstonia s
36	54	19.9	488	1	PAC2_RAT	Q9qy17 rattus norv
37	54	19.9	1085	1	CUT7_SCHPO	P24339 schizosacch
38	54	19.9	2280	1	COAC_SCHPO	P78820 schizosacch
39	53.5	19.7	199	1	PEXH_YEAST	P40155 saccharomyc
40	53.5	19.7	270	1	TCMP_STRGA	P39887 streptomyc
41	53.5	19.7	323	1	VALL_PASVK	Q00338 panicum str
42	53.5	19.7	462	1	MYCN_MOUSE	P03966 mus musculu
43	53.5	19.7	1092	1	CARB_ZYMMO	O50236 zymomonas m
44	53	19.6	288	1	MRGX_HUMAN	Q15014 homo sapien
45	53	19.6	772	1	IMBT_HUMAN	Q9y468 homo sapien

ALIGNMENTS

RESULT 1					
CPSM_HUMAN					
ID	CPSM_HUMAN	STANDARD;	PRT;	1500	AA.
AC	P31327; O43774;				
DT	01-JUL-1993 (Rel. 26, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Carbamoyl-phosphate synthase [ammonia], mitochondrial precursor				
DE	(EC 6.3.4.16) (Carbamoyl-phosphate synthetase I) (CPSase I).				
GN	CPS1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Liver;				
RX	MEDLINE=92084128; PubMed=1840546;				
RA	Haraguchi Y., Uchino T., Takiguchi M., Endo F., Mori M.,				
RA	Matsuda I.;				
RT	"Cloning and sequence of a cDNA encoding human carbamyl phosphate				
RT	synthetase I: molecular analysis of hyperammonemia.";				
RL	Gene 107:335-340(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A., VARIANT CPS1 DEFICIENCY MET-544, AND VARIANT				
RP	ALA-344.				
RC	TISSUE=Liver;				
RX	MEDLINE=98375696; PubMed=9711878;				
RA	Finckh U., Kohlschuetter A., Schaefer H., Sperhake K., Colombo J.-P.,				
RA	Gal A.;				
RT	"Prenatal diagnosis of carbamoyl phosphate synthetase I deficiency by				
RT	identification of a missense mutation in CPS1.";				
RL	Hum. Mutat. 12:206-211(1998).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RA	Summar M.;				
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	VARIANT CPS1 DEFICIENCY ARG-337.				
RX	MEDLINE=21367742; PubMed=11474210;				
RA	Aoshima T., Kajita M., Sekido Y., Kikuchi S., Yasuda I., Saheki T.,				
RA	Watanabe K., Shimokata K., Niwa T.;				
RT	"Novel mutations (H337R and 238-362del) in the CPS1 gene cause				
RT	carbamoyl phosphate synthetase I deficiency.";				
RL	Hum. Hered. 52:99-101(2001).				
CC	!- FUNCTION: INVOLVED IN THE UREA CYCLE OF UREOTELIC ANIMALS WHERE				
CC	THE ENZYME PLAYS AN IMPORTANT ROLE IN REMOVING EXCESS AMMONIA				
CC	FROM THE CELL.				
CC	!- CATALYTIC ACTIVITY: 2 ATP + NH(3) + CO(2) + H(2)O = 2 ADP +				
CC	phosphate + carbamoyl phosphate.				
CC	!- ENZYME REGULATION: REQUIRES N-ACETYLGLUTAMATE AS AN ALLOSTERIC				
CC	ACTIVATOR.				
CC	!- SUBCELLULAR LOCATION: Mitochondrial.				
CC	!- TISSUE SPECIFICITY: PRIMARILY IN THE LIVER AND SMALL INTESTINE.				
CC	!- DISEASE: Defects in CPS1 are the cause of an autosomal recessive				
CC	metabolic disorder that cause a type of hyperammonemia. Clinical				

A;Reference number: S62445
A;Accession: S62450
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 'MSGLLPSLSSSPFLVQSEALG',1-314,'GIC',318-1013,'CAVR',1017,1019-1387,'EL',1390
A;Cross-references: EMBL:Z54328; NID:g1009451; PIDN:CAA91130.1; PID:g1009456
C;Genetics:
A;Gene: URA1
A;Map position: 1L
C;Superfamily: rudimentary enzyme; aspartate/ornithine carbamoyltransferase homology; Ba
arbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homology; carbamoyl-phos
C;Keywords: ligase; multifunctional enzyme; pyrimidine nucleotide biosynthesis; transfer
F;40-1498/Domain: carbamoyl-phosphate synthase (ammonia) homology <CPA>
F;40-419/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain homolo
F;244-419/Domain: trpG homology <TRG>
F;455-1495/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain hom
F;455-905/Domain: biotin carboxylase homology <BC1>
F;994-1437/Domain: biotin carboxylase homology <BC2>
F;1511-1836/Domain: Bacillus dihydroorotase homology <DHO>
F;1916-2216/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>

Query Match 23.2%; Score 63; DB 2; Length 2223;
Best Local Similarity 35.6%; Pred. No. 32;
Matches 16; Conservative 10; Mismatches 17; Indels 2; Gaps 2;

QY 1 QLNHEGFKLFATEATSDWLNANNVPATPVA-WPSQEGQNP-SLSS 43
:|: :||| ||| :|| :| |::| |||:
Db 1395 KLYENNYNIFATAGTSDYFMEGVPCKYLADLPAAEEANNEYSLSA 1439

Search completed: September 13, 2003, 02:16:43
Job time : 24.5 secs

R;Denis-Duphil, M.; Lecaer, J.P.; Hardie, D.G.; Carrey, E.A.
Eur. J. Biochem. 193, 581-587, 1990
A;Title: Yeast carbamoyl-phosphate-synthetase - aspartate-transcarbamylase multidomain p
A;Reference number: S13358; MUID:91031508; PMID:1977585
A;Accession: S13358
A;Molecule type: protein
A;Residues: 1855-1882 <DEN>
R;Katsoulou, C.; Tzernia, M.; Tavernarakis, N.; Alexandraki, D.
Yeast 12, 787-797, 1996
A;Title: Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome X
of Chromosome XI.
A;Reference number: S71643; MUID:96408771; PMID:8813765
A;Accession: S71667
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 175-2214 <KAW>
A;Cross-references: EMBL:X87371; NID:9854542; PIDN:CAA60825.1; PID:9854566
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
R;Nagy, M.; le Gouar, M.; Potier, S.; Souciet, J.L.; Herve, G.
J. Biol. Chem. 264, 8366-8374, 1989
A;Title: The primary structure of the aspartate transcarbamylase region of the URA2 gene
A;Reference number: A33820; MUID:89255278; PMID:2498313
A;Accession: S22790
A;Molecule type: DNA
A;Residues: 1268-1410, 'S', 1413-1581, 'M', 1583-1587, 'K', 1589-1591, 'G', 1593-1594, 'A', 1596-2
A;Cross-references: EMBL:J04711
C;Genetics:
A;Gene: SGD:URA2; MIPS:YJL130C
A;Cross-references: SGD:S0003666; MIPS:YJL130C
A;Map position: 10L
C;Function:
A;Description: ligase; methyltransferase
A;Pathway: pyrimidine nucleotide biosynthesis
C;Superfamily: rudimentary enzyme; aspartate/ornithine carbamoyltransferase homology; Ba
arbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homology; carbamoyl-phos
C;Keywords: ligase; methyltransferase; multifunctional enzyme; nucleus; phosphoprotein;
F;24-1487/Domain: carbamoyl-phosphate synthase (ammonia) homology <CPA>
F;24-404/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain homolo
F;229-404/Domain: trpG homology <TRG>
F;440-1484/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homo
F;440-891/Domain: biotin carboxylase homology <BC1>
F;981-1424/Domain: biotin carboxylase homology <BC2>
F;1500-1825/Domain: Bacillus dihydroorotase homology <DHO>
F;1911-2210/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>
F;302/Active site: Cys #status predicted

Query Match 24.0%; Score 65; DB 1; Length 2214;
Best Local Similarity 45.8%; Pred. No. 18;
Matches 11; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 QLNHEGKLFATEATSDWLNANNV 24
Db 1382 KLYNMGYKLFATSGTADFLSEHGI 1405

RESULT 13
AE2722
DNA-directed RNA polymerase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AE2722
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyaev, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AE2722
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-863 <KUR>

A;Cross-references: GB:AE008688; PIDN:AAL42195.1; PID:gl7739586; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atul183
A;Map position: circular chromosome
C;Superfamily: phage T7 DNA-directed RNA polymerase

Query Match 23.6%; Score 64; DB 2; Length 863;
Best Local Similarity 30.2%; Pred. No. 7.6;
Matches 19; Conservative 7; Mismatches 23; Indels 14; Gaps 2;

QY 3 HNEGFKLFATEATSDWL---NANNVPATPVAMPQEQGNPSLSIR-----KLI 48
Db 221 HNTNIYVATKETLDWLATENSRLAPLSPVILPTLVPPRPWTSPFRGGYWSGVRNLRLLI 280

QY 49 RDG 51
Db 281 KTG 283

RESULT 14
B97504

RNA polymerase [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: B97504
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: B97504
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-863 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK86987.1; PID:gl5156227; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C_2186
A;Map position: circular chromosome
C;Superfamily: phage T7 DNA-directed RNA polymerase

Query Match 23.6%; Score 64; DB 2; Length 863;
Best Local Similarity 30.2%; Pred. No. 7.6;
Matches 19; Conservative 7; Mismatches 23; Indels 14; Gaps 2;

QY 3 HNEGFKLFATEATSDWL---NANNVPATPVAMPQEQGNPSLSIR-----KLI 48
Db 221 HNTNIYVATKETLDWLATENSRLAPLSPVILPTLVPPRPWTSPFRGGYWSGVRNLRLLI 280

QY 49 RDG 51
Db 281 KTG 283

RESULT 15
S65074

pyrimidine synthesis protein URAL - fission yeast (Schizosaccharomyces pombe)
N;Alternate names: protein SPAC22G7.06c
N;Contains: aspartate carbamoyltransferase (EC 2.1.3.2); carbamoyl-phosphate synthase
C;Species: Schizosaccharomyces pombe
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C;Accession: S65074; S62450
R;Lollier, M.; Jaquet, L.; Nedeva, T.; Lacroute, F.; Potier, S.; Souciet, J.L.
Curr. Genet. 28, 138-149, 1995
A;Title: As in Saccharomyces cerevisiae, aspartate transcarbamoylase is assembled on
A;Reference number: S65074; MUID:96020160; PMID:8590465
A;Accession: S65074
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2223 <LOL>
A;Cross-references: EMBL:X81841; NID:gl154951; PIDN:CAA57433.1; PID:gl154952
R;Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, October 1995


```

Matches 13; Conservative 9; Mismatches 11; Indels 1; Gaps 1;
QY      6 GFKLFAATEATSDWLNANNVPATPVAWPSQE-QQN 38
       I : | | | : | : | | | : | : | : | : |
Db      950 GTTLFGTQTADFYSENGVPTQLNWDEEDLGEN 983

RESULT 8
QZDOP3
pyrimidine synthesis protein PYR1-3 - slime mold (Dictyostelium discoideum) (fragment
N;Contains: aspartate carbamoyltransferase (EC 2.1.3.2); carbamoyl-phosphate synthase
C;Species: Dictyostelium discoideum
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 07-Aug-1998
C;Accession: S02800
R;Faure, M.; Camonis, J.H.; Jacquet, M.
Eur. J. Biochem. 179, 345-358, 1989
A;Title: Molecular characterization of a Dictyostelium discoideum gene encoding a mul
A;Reference number: S02800; MUID:89137111; PMID:2917570
A;Accession: S02800
A;Molecule type: DNA
A;Residues: 1-457;468-1481 <FAU>
A;Cross-references: EMBL:X14633
A;Experimental source: strain AX3
C;Genetics:
A;Gene: PYR1-3
C;Superfamily: rudimentary enzyme; aspartate/ornithine carbamyltransferase homology;
arbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homology; carbamoyl-p
C;Keywords: hydrolase; ligase; methyltransferase; multifunctional enzyme; pyrimidine
F:1-707/Domain: carbamoyl-phosphate synthase (ammonia) homology (fragments) <CPA>
F:1-340/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain homo
F:158-340/Domain: trpg homology <TRG>
F:357-704/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain ho
F:367-467/Domain: biotin carboxylase homology (fragment) <BC1>
F:468-651/Domain: biotin carboxylase homology (fragment) <BC2>
F:721-1066/Domain: Bacillus dihydroorotase homology <DHO>
F:1179-1477/Domain: aspartate/ornithine carbamyltransferase homology <ACT>
F:236/Active site: Cys #status predicted

Query Match          27.5%; Score 74.5; DB 1; Length 1481;
Best Local Similarity 38.2%; Pred. No. 0.68;
Matches 13; Conservative 9; Mismatches 11; Indels 1; Gaps 1;
QY      6 GFKLFAATEATSDWLNANNVPATPVAWPSQE-QQN 38
       I : | | | : | : | | | : | : | : | : |
Db      611 GTTLFGTQTADFYSENGVPTQLNWDEEDLGEN 644

RESULT 9
B83605
probable acetylpolymaine aminohydrolase PA0321 [imported] - Pseudomonas aeruginosa (s
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: B83605
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: B83605
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-344 <STO>
A;Cross-references: GB:AE004470; GB:AE004091; NID:g9946164; PIDN:AAG03710.1; GSPDB:GN
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA0321

Query Match          25.6%; Score 69.5; DB 2; Length 344;
Best Local Similarity 37.1%; Pred. No. 0.5;
Matches 13; Conservative 7; Mismatches 14; Indels 1; Gaps 1;
QY      1 QLNHEGFKLFATEATSDWL-NANNVPATPVAWPSQ 34

```

```

RESULT 2
JQ1348
carbamoyl-phosphate synthase (ammonia) (EC 6.3.4.16) precursor - human
N;Alternate names: carbamyl phosphate synthetase I; carbon-dioxide-ammonia ligase
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1992 #sequence_revision 02-Jun-1994 #text_change 16-Jun-2000
C;Accession: JQ1348
R;Haraguchi, Y.; Uchino, T.; Takiguchi, M.; Endo, F.; Mori, M.; Matsuda, I.
Gene 107, 335-340, 1991
A;Title: Cloning and sequence of a cDNA encoding human carbamyl phosphate synthetase I:
A;Reference number: JQ1348; MUID:92084128; PMID:1840546
A;Accession: JQ1348
A;Molecule type: mRNA
A;Residues: 1-1500 <HAR>
A;Cross-references: DDBJ:D90282; NID:g219552; PIDN:BAA14328.1; PID:g219553
C;Comment: This is the first enzyme of the urea cycle; it catalyzes the synthesis of car
C;Genetics:
A;Gene: GDB:CPS1
A;Cross-references: GDB:119799; OMIM:237300
A;Map position: 2q33-2q36
C;Superfamily: carbamoyl-phosphate synthase (ammonia); biotin carboxylase homology; carb
omology; carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain homology; trpG
C;Keywords: ATP; ligase; urea cycle
F;1-38/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;39-1500/Product: carbamoyl-phosphate synthase (ammonia) #status predicted <CAR>
F;48-1479/Domain: carbamoyl-phosphate synthase (ammonia) homology <CPA>
F;48-395/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain homolo
F;220-395/Domain: trpG homology <TRG>
F;424-1476/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain hom
F;424-886/Domain: biotin carboxylase homology <BC1>
F;975-1426/Domain: biotin carboxylase homology <BC2>

Query Match 98.2%; Score 266; DB 1; Length 1500;
Best Local Similarity 98.0%; Pred. No. 3.1e-25;
Matches 50; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QLNHEGFKLFATEATSDWLNANNVPATPVWPSQEQNPSSLIRKILRDG 51
|||||
Db 1380 QLNHEGFKLFATEATSDWLNANNVPANPVWPSQEQNPSSLIRKILRDG 1430

RESULT 3
I51170
carbamyl phosphate synthetase - bullfrog
C;Species: Rana catesbeiana (bullfrog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 22-Jun-1999
C;Accession: I51170
R;Helbing, C.C.; Atkinson, B.G.
J. Biol. Chem. 269, 11743-11750, 1994
A;Title: 3,5,3'-Triiodothyronine-induced carbamyl phosphate synthetase gene expression in
A;Reference number: A53567; MUID:94216272; PMID:8163471
A;Accession: I51170
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1496 <HEL>
A;Cross-references: EMBL:U05193; NID:g467222; PIDN:AAAL9016.1; PID:g467223
C;Superfamily: carbamoyl-phosphate synthase (ammonia); biotin carboxylase homology; carb
omology; carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain homology; trpG
C;Keywords: ATP
F;43-1475/Domain: carbamoyl-phosphate synthase (ammonia) homology <CPA>
F;43-392/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain homolo
F;216-392/Domain: trpG homology <TRG>
F;421-1472/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain hom
F;421-883/Domain: biotin carboxylase homology <BC1>
F;972-1422/Domain: biotin carboxylase homology <BC2>
F;290/Active site: Cys #status predicted

```

QY 2 LHNEGFKLFATEATSDWLNANNVPATPVAFWSQEGQNPSLSIRKLIRDG 51
I :|||||:|||||:|||||:|||||: |||||:
Db 1378 LKDEGFKLYATEATADWLNANDITATPVAFWSQEQS-GPSSIIYKLIKEG 1426

RESULT 4
S53602 carbamoyl-phosphate synthase (glutamine-hydrolyzing) (EC 6.3.5.5) III precursor, hepa
N;Alternate names: carbamyl phosphate synthetase III
C;Species: Squalus acanthias (spiny dogfish)
C;Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 07-Dec-1999
C;Accession: S53602; S72160; S71062
R;Hong, J.; Salo, W.L.; Lusty, C.J.; Anderson, P.M.
J. Mol. Biol. 243, 131-140, 1994
A;Title: Carbamyl phosphate synthetase III, an evolutionary intermediate in the trans
A;Reference number: S53602; MUID:95018247; PMID:7932737
A;Accession: S53602
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1502 <HON1>
A;Cross-references: EMBL:L31362
A;Accession: S72160
A;Molecule type: protein
A;Residues: 39-47 <HON2>
R;Anderson, P.M.
submitted to the EMBL Data Library, September 1994
A;Reference number: S71062
A;Accession: S71062
A;Molecule type: mRNA
A;Residues: 1-888, E', 890-1502 <AND>
A;Cross-references: EMBL:L31362; NID:g1256714; PIDN:AAA96435.1; PID:g530209
C;Genetics:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: September 13, 2003, 02:11:17 ; Search time 23.5 Seconds
(without alignments)
208.707 Million cell updates/sec

Title: US-09-585-077C-4_COPY_1380_1430
Perfect score: 271
Sequence: 1 QLHNEGFKLFATEATSDWLN.....PSQEQGNPSLSIRKLIRDG 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	271	100.0	1500	1 SYRTCA	carbamoyl-phosphat
2	266	98.2	1500	1 JQ1348	carbamoyl-phosphat
3	184.5	68.1	1496	2 I51170	carbamyl phosphate
4	165	60.9	1502	2 S53602	carbamoyl-phosphat
5	82	30.3	618	2 D64472	carbamoyl-phosphat
6	79	29.2	2198	2 T20371	hypothetical prote
7	74.5	27.5	1042	2 S23738	pyrl-3 protein - s
8	74.5	27.5	1481	1 QZDOP3	pyrimidine synthes
9	69.5	25.6	344	2 B83605	probable acetylpol
10	69.5	25.6	1115	2 A70990	carbamoyl-phosphat
11	68.5	25.3	1071	2 F39845	carbamoyl-phosphat
12	65	24.0	2214	1 QZBYU2	pyrimidine synthes
13	64	23.6	863	2 AE2722	DNA-directed RNA p
14	64	23.6	863	2 B97504	RNA polymerase [im
15	63	23.2	2223	2 S65074	pyrimidine synthes
16	63	23.2	2244	2 T11616	carbamoyl-phosphat
17	62	22.9	336	2 G95003	membrane protein [
18	62	22.9	340	2 B97876	conserved hypothet
19	59.5	22.0	537	2 B87414	RNA polymerase sig
20	59.5	22.0	1249	2 T14150	vesicle associated
21	59.5	22.0	2225	1 A23443	pyrimidine synthes
22	58.5	21.6	277	2 AE1182	sugar ABC transpor
23	58.5	21.6	277	2 AF1539	sugar ABC transpor
24	58.5	21.6	715	2 S54628	hypothetical prote
25	58	21.4	452	2 C72277	phosphoribosylamin
26	57	21.0	2242	2 A57541	pyrimidine synthes
27	56.5	20.8	673	4 F40201	artifact-warning s
28	56.5	20.8	700	2 B82788	metallopeptidase x
29	56.5	20.8	771	2 S51421	hypothetical prote

probable carbamoyl
hypothetical prote
carbamoyl-phosphat
hypothetical prote
hypothetical prote
hypothetical prote
vegetative storage
vegetative storage
vegetative storage
probable lipoprote
probable serine-ty
probable membrane
carbamoyl-phosphat
surface layer prot
hypothetical prote
hypothetical prote

30 56.5 20.8 1129 2 H86975
31 56 20.7 297 2 T18637
32 56 20.7 1162 2 AD3317
33 56 20.7 1524 2 G84721
34 56 20.7 1750 2 G84649
35 55 20.3 190 2 G84018
36 55 20.3 244 1 UESY25
37 55 20.3 254 2 S08511
38 55 20.3 291 2 T08848
39 55 20.3 521 2 S62794
40 55 20.3 2105 2 T18968
41 54.5 20.1 662 2 S56303
42 54.5 20.1 1062 2 H83966
43 54.5 20.1 1228 2 I40468
44 54 19.9 347 2 S55629
45 54 19.9 781 2 T41551

ALIGNMENTS

RESULT 1

SYRTCA

carbamoyl-phosphate synthase (ammonia) (EC 6.3.4.16) I precursor - rat
N;Alternate names: carbamyl phosphate synthetase I; carbon-dioxide-ammonia ligase
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 31-Dec-2000
C;Accession: A28481; A23580
R;Lagace, M.; Howell, B.W.; Burak, R.; Lusty, C.J.; Shore, G.C.
J. Biol. Chem. 262, 10415-10418, 1987
A;Title: Rat carbamyl-phosphate synthetase I gene. Promoter sequence and tissue-speci
A;Reference number: A28481; MUID:87280088; PMID:3038878
A;Accession: A28481
A;Molecule type: DNA
A;Residues: 1-42 <LAG>
A;Cross-references: GB:J02805; NID:G203577; PIDN:AAA40959.1; PID:G203578
R;Nyunoya, H.; Broglie, K.E.; Widgren, E.E.; Lusty, C.J.
J. Biol. Chem. 260, 9346-9356, 1985
A;Title: Characterization and derivation of the gene coding for mitochondrial carbamy
A;Reference number: A23580; MUID:85261323; PMID:2991241
A;Accession: A23580
A;Molecule type: mRNA
A;Residues: 1-1500 <NYU>
A;Cross-references: GB:M11710
R;Potter, M.D.; Powers-Lee, S.G.
J. Biol. Chem. 267, 2023-2031, 1992
A;Title: Location of the ATP-gamma-phosphate-binding sites on rat liver carbamoyl-pho
A;Reference number: A42051; MUID:92112936; PMID:1730733
A;Contents: annotation; MgATP binding sites
C;Comment: This enzyme, which requires N-acetylglutamate as allosteric activator, cat
C;Comment: This enzyme of the mitochondrial matrix is expressed specifically in hepat
C;Comment: This enzyme binds two molecules of ATP at discrete sites for use in discr
C;Superfamily: carbamoyl-phosphate synthase (ammonia); biotin carboxylase homology; c
omology; carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain homology; t
C;Keywords: ATP; duplication; ligase; mitochondrion; urea cycle
F;1-38/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;39-1500/Product: carbamoyl-phosphate synthase (ammonia) #status predicted <MAT>
F;48-1479/Domain: carbamoyl-phosphate synthase (ammonia) homology <CPA>
F;48-395/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain hom
F;220-395/Domain: trpG homology <TRG>
F;424-1476/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain h
F;424-886/Domain: biotin carboxylase homology <BCI>
F;631-638,1327-1348/Region: ATP binding (for bicarbonate activation) #status experime
F;975-1426/Domain: biotin carboxylase homology <BC2>
F;1310-1317,1445-1454/Region: ATP binding (for carbamate activation) #status experime

Query Match 100.0%; Score 271; DB 1; Length 1500;
Best Local Similarity 100.0%; Pred. No. 7.2e-26;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLHNEGFKLFATEATSDWLNANNVPATPVAPWSQEQGNPSLSIRKLIRDG 51

Db 1380 QLHNEGFKLFATEATSDWLNANNVPATPVAPWSQEQGNPSLSIRKLIRDG 1430

Search completed: September 13, 2003, 02:12:45
Job time : 53 secs

QY
4 NEGKLFATEATSDWLNANNVPATPVAWPSQGQNPSLSSI 44
- - - - : : : : : : : : : :
db 21 NSPFKFTGRGOAOWL---MPVTPALWEAKVGCRSPVRSL 57

1 OLNHNEGKLFATEATSDWLNANNVPATPVAWPSEOEGONPSLSSIRKLIRD 50

CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC disorders involving aberrant protein expression or biological activity.
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 692 AA;

Query Match 23.4%; Score 63.5; DB 22; Length 692;
Best Local Similarity 37.8%; Pred. No. 9.7;
Matches 14; Conservative 7; Mismatches 9; Indels 7; Gaps 2;

QY 1 QLHNEGFKLFATEATSDWLNANNVPATPPVAMPVSQEGQ 37
:||::||||: | |:: |::| | |||
Db 232 ELHHKGFKLWFDELIGLWVSVQNLP-----QW---EGQ 261

RESULT 13
ABG09933
ID ABG09933 standard; Protein; 1798 AA.
XX
AC ABG09933;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #9924.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS74120.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 40292; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1798 AA;

Query Match 23.4%; Score 63.5; DB 22; Length 1798;
Best Local Similarity 37.8%; Pred. No. 34;
Matches 14; Conservative 7; Mismatches 9; Indels 7; Gaps 2;

QY 1 QLHNEGFKLFATEATSDWLNANNVPATPPVAMPVSQEGQ 37
:||::||||: | |:: |::| | |||
Db 180 ELHHKGFKLWFDELIGLWVSVQNLP-----QW---EGQ 209

RESULT 14
AAO11745
ID AAO11745 standard; Protein; 60 AA.
XX
AC AAO11745;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 25637.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI91676.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 20; SEQ ID NO 25637; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and

Best Local Similarity 37.0%; Pred. No. 0.31;
Matches 17; Conservative 1; Mismatches 12; Indels 16; Gaps 2

QY 18 WLNNANVPATPVAW-----PSQEGQNPSLSSIRKLIRDG 51
II I III I
DB 13 WL----TPVIPVLWDTKADSGIRDQPQSQHGETPSLLKIEKLAGHG 54

RESULT 9
ABP75436
ID ABP75436 standard; Protein; 94 AA.
XX AC ABP75436;
XX DT 10-FEB-2003 (first entry)
XX DE Human secretory polypeptide SPTM SEQ ID NO 620.
XX KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;
KW anti-inflammatory; immunosuppressive; neuroprotective; nootropic;
KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;
KW antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;
KW secretory polynucleotide; secretory protein.

XX OS Homo sapiens.
XX PN WO200283876-A2.
XX PD 24-OCT-2002.
XX PF 27-MAR-2002; 2002WO-US09921.
XX PR 29-MAR-2001; 2001US-280067P.
PR 29-MAR-2001; 2001US-280068P.
PR 16-MAY-2001; 2001US-291280P.
PR 17-MAY-2001; 2001US-291829P.
PR 17-MAY-2001; 2001US-291849P.
PR 19-JUN-2001; 2001US-299428P.
PR 20-JUN-2001; 2001US-299776P.
PR 20-JUN-2001; 2001US-300001P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX WPI; 2003-075543/07.
DR N-PSDB; ABZ35886.
XX PT New human secretory proteins and polynucleotides, useful for
PT diagnosing, treating or preventing autoimmune/inflammatory disorders
PT (e.g. AIDS), neurological disorders (e.g. Alzheimer's), or cell
PT proliferations or cancers -
XX PS Claim 27; SEQ ID NO 620; 458pp + Sequence Listing; English.
XX CC The invention relates to a secretory polynucleotide (designated sptm)
CC comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a
CC naturally occurring polynucleotide sequence at least 90 % identical to
CC the polynucleotide sequence, a polynucleotide complementary to them or an
CC RNA equivalent of them. The polypeptide or polynucleotide are useful for
CC treating, preventing or diagnosing a disease or condition associated with
CC the expression of functional SPTM. These are particularly useful for
CC diagnosing, treating or preventing autoimmune/inflammatory disorders
CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's
CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,
CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,

multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety, schizophrenia or amnesia), or cell proliferative disorders (e.g. psoriasis, polycythemia vera, or cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma; sarcoma or cancers of the brain, breast, cervix or prostate). The present sequence is one of the SPTM proteins of the invention (ABP75384-ABP75962).

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Query Match	23.8%;	Score 64.5;	DB 24;	Length 94;
Best Local Similarity	42.4%;	Pred. No. 0.51;		
Matches 14;	Conservative	3;	Mismatches 13;	Indels 3;
				Gaps 17;

QY 13 EATSDWLNANVPATPVA---WPSQEGQNPSLS 42
| : | : | | | | | | | : |
Db 33 ETADWSDAATSPGTPGAPSWEROEGSSPRAS 65

RESULT 10
AAU03138
ID AAU03138 standard; Protein: 339 AA;

Query Match	23.8%;	Score 64.5;	DB 24;	Length 94;
Best Local Similarity	42.4%;	Pred. No. 0.51;		
Matches 14;	Conservative	3;	Mismatches 13;	Indels 3

OS Homo sapiens.
XX WO200073322-A1.
PN
XX 07-DEC-2000.
PD
XX 01-JUN-2000; 2000WO-US15079.
PF
XX 01-JUN-1999; 99US-0323472.
PR
XX (UYVA-) UNIV VANDERBILT.
XX Summar ML, Christman BW;
PI
XX WPI; 2001-049926/06.
DR
XX Detecting susceptibility to suboptimal urea cycle function, e.g. bone
PT marrow transplant toxicity, comprises identifying a polymorphism in the
PT gene for carbamylphosphate synthase -
XX
PS Claim 57; Page 152-155; 171pp; English.
XX
CC The present invention relates to screening for susceptibility to
CC suboptimal urea cycle function or to bone marrow transplant toxicity
CC by detecting a polymorphism in the carbamyl phosphate synthase I
CC (CPSI) gene. The method is used to detect subjects at risk of
CC hepatitis, sclerosis, pulmonary hypertension and bone marrow
CC transplant toxicity. These conditions can be treated or prevented
CC by administration of a nitric oxide precursor or by gene
CC therapy (administration of sequences that encode CPSI).
XX
XX Sequence 1500 AA;
SQ

Query Match 98.2%; Score 266; DB 22; Length 1500;
Best Local Similarity 98.0%; Pred. No. 1.7e-27;
Matches 50; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLNHEGKLFATEATSDWLNANNVPATPVAWPSQEGQNPSLSIRKLIRDG 51
Db 1380 QLNHEGKLFATEATSDWLNANNVPANPVAWPSQEGQNPSLSIRKLIRDG 1430

RESULT 5
AAB49225
ID AAB49225 standard; protein; 1500 AA.
XX
AC AAB49225;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human carbamyl phosphate synthase protein.
XX
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy.
XX
OS Homo sapiens.
OS
XX
XX WO200073322-A1.
PN
XX 07-DEC-2000.
PD
XX
PF 01-JUN-2000; 2000WO-US15079.
PR
XX 01-JUN-1999; 99US-0323472.
XX
XX (UYVA-) UNIV VANDERBILT.
PA
XX Summar ML, Christman BW;
PI
XX WPI; 2001-049926/06.
DR
XX Detecting susceptibility to suboptimal urea cycle function, e.g. bone
PT marrow transplant toxicity, comprises identifying a polymorphism in the
PT

PT gene for carbamylphosphate synthase -
XX
PS Claim 57; Page 162-165; 171pp; English.
XX
CC The present invention relates to screening for susceptibility to
CC suboptimal urea cycle function or to bone marrow transplant toxicity
CC by detecting a polymorphism in the carbamyl phosphate synthase I
CC (CPSI) gene. The method is used to detect subjects at risk of
CC hepatitis, sclerosis, pulmonary hypertension and bone marrow
CC transplant toxicity. These conditions can be treated or prevented
CC by administration of a nitric oxide precursor or by gene
CC therapy (administration of sequences that encode CPSI).
XX
XX Sequence 1500 AA;
SQ

Query Match 98.2%; Score 266; DB 22; Length 1500;
Best Local Similarity 98.0%; Pred. No. 1.7e-27;
Matches 50; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLNHEGKLFATEATSDWLNANNVPATPVAWPSQEGQNPSLSIRKLIRDG 51
Db 1380 QLNHEGKLFATEATSDWLNANNVPANPVAWPSQEGQNPSLSIRKLIRDG 1430

RESULT 6
AAU33563
ID AAU33563 standard; Protein; 344 AA.
XX
AC AAU33563;
XX
DT 14-FEB-2002 (first entry)
XX
DE Pseudomonas aeruginosa cellular proliferation protein #7.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Pseudomonas aeruginosa.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS51422.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 5059; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets

XX PS Disclosure; Fig 11; 171pp; English.

XX CC The present invention relates to screening for susceptibility to

CC suboptimal urea cycle function or to bone marrow transplant toxicity

CC by detecting a polymorphism in the carbamyl phosphate synthase I

CC (CPSI) gene. The method is used to detect subjects at risk of

CC hepatitis, sclerosis, pulmonary hypertension and bone marrow

CC transplant toxicity. These conditions can be treated or prevented

CC by administration of a nitric oxide precursor or by gene

CC therapy (administration of sequences that encode CPSI).

XX SQ Sequence 1500 AA;

Query Match 100.0%; Score 271; DB 22; Length 1500;

Best Local Similarity 100.0%; Pred. No. 3.4e-28;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLHNEGFKLFATEATSDWLNANNVPATPVAVPWSQEQNPSSLSSIRKLIRDG 51

Db 1380 QLHNEGFKLFATEATSDWLNANNVPATPVAVPWSQEQNPSSLSSIRKLIRDG 1430

RESULT 2

AAR30636

ID AAR30636 standard; Protein; 1500 AA.

XX AC AAR30636;

XX DT 06-MAY-1993 (first entry)

XX DE hCPSI.

XX KW Human; carbamyl phosphate synthetase I; CPSI; probe; mutation;

XX KW CPSI deficiency disease; rat.

XX OS Homo sapiens.

XX PN JP04335889-A.

XX PD 24-NOV-1992.

XX PF 09-MAY-1991; 91JP-0135902.

XX PR 09-MAY-1991; 91JP-0135902.

XX (HARA/) HARAGUCHI Y.

PA (MATS/) MATSUDA I.

PA (MORI/) MORI M.

XX WPI; 1993-006237/01.

DR N-PSDB; AAQ34768.

XX Carbamyl phosphate synthetase I gene - used to detect diseases

PT caused by carbamyl phosphatase synthetase by using overlapped

PT cDNA contg. total human cDNA sequence

XX Disclosure; Fig 2-9; 12pp; Japanese.

XX This sequence is the human carbamyl phosphate synthetase I (CPSI)

CC protein. The DNA encoding this peptide or fragments of it may be

CC used as probes to detect mutations in the CPSI gene. CPSI deficiency

CC disease may be detected by using overlapping cDNA representing the

CC full length cDNA sequence of human CPSI. The DNA sequence was

CC isolated using three amplified fragments derived from protein coding

CC regions of the rat CPSI cDNA as probes.

XX SQ Sequence 1500 AA;

Query Match 98.2%; Score 266; DB 14; Length 1500;

Best Local Similarity 98.0%; Pred. No. 1.7e-27;

Matches 50; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLHNEGFKLFATEATSDWLNANNVPATPVAVPWSQEQNPSSLSSIRKLIRDG 51

Db 1380 QLHNEGFKLFATEATSDWLNANNVPATPVAVPWSQEQNPSSLSSIRKLIRDG 1430

RESULT 4

AAB49224

ID AAB49224 standard; protein; 1500 AA.

XX AC AAB49224;

XX DT 13-MAR-2001 (first entry)

XX DE Human carbamyl phosphate synthase protein.

XX KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;

XX KW sclerosis; pulmonary hypertension; bone marrow; gene therapy.

QY 1 QLHNEGFKLFATEATSDWLNANNVPATPVAVPWSQEQNPSSLSSIRKLIRDG 51

Db 1380 QLHNEGFKLFATEATSDWLNANNVPANPVAVPWSQEQNPSSLSSIRKLIRDG 1430

RESULT 3

AAB49222

ID AAB49222 standard; protein; 1500 AA.

XX AC AAB49222;

XX DT 13-MAR-2001 (first entry)

XX DE Human N1405 carbamyl phosphate synthase I protein.

XX KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;

XX KW sclerosis; pulmonary hypertension; bone marrow; gene therapy.

XX OS Homo sapiens.

XX PN WO200073322-A1.

XX PD 07-DEC-2000.

XX PF 01-JUN-2000; 2000WO-US15079.

XX PR 01-JUN-1999; 99US-0323472.

XX PA (UYVA-) UNIV VANDERBILT.

PI Summar ML, Christman BW;

XX WPI; 2001-049926/06.

DR Detecting susceptibility to suboptimal urea cycle function, e.g. bone

XX marrow transplant toxicity, comprises identifying a polymorphism in the

PT gene for carbamylphosphate synthase -

XX Claim 57; Fig 12; 171pp; English.

XX The present invention relates to screening for susceptibility to

CC suboptimal urea cycle function or to bone marrow transplant toxicity

CC by detecting a polymorphism in the carbamyl phosphate synthase I

CC (CPSI) gene. The method is used to detect subjects at risk of

CC hepatitis, sclerosis, pulmonary hypertension and bone marrow

CC transplant toxicity. These conditions can be treated or prevented

CC by administration of a nitric oxide precursor or by gene

CC therapy (administration of sequences that encode CPSI).

XX SQ Sequence 1500 AA;

Query Match 98.2%; Score 266; DB 22; Length 1500;

Best Local Similarity 98.0%; Pred. No. 1.7e-27;

Matches 50; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLHNEGFKLFATEATSDWLNANNVPATPVAVPWSQEQNPSSLSSIRKLIRDG 51

Db 1380 QLHNEGFKLFATEATSDWLNANNVPANPVAVPWSQEQNPSSLSSIRKLIRDG 1430

RESULT 4

AAB49224

ID AAB49224 standard; protein; 1500 AA.

XX AC AAB49224;

XX DT 13-MAR-2001 (first entry)

XX DE Human carbamyl phosphate synthase protein.

XX KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;

XX KW sclerosis; pulmonary hypertension; bone marrow; gene therapy.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2003, 00:29:37 ; Search time 52 Seconds
(without alignments)
155.674 Million cell updates/sec

Title: US-09-585-077C-4_COPY_1380_1430
Perfect score: 271
Sequence: 1 QLNHEGFKLFATEATSDWLN.....PSQEQNPSSLIRKLIRDG 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	271	100.0	1500	22 AAB49223	Human T1405 carbam
2	266	98.2	1500	14 AAR30636	hCPSI. Homo sapie
3	266	98.2	1500	22 AAB49222	Human N1405 carbam
4	266	98.2	1500	22 AAB49224	Human carbamyl pho
5	266	98.2	1500	22 AAB49225	Human carbamyl pho
6	69.5	25.6	344	22 AAU33563	Pseudomonas aerugi
7	68	25.1	43	22 AAO13747	Human polypeptide
8	66	24.4	94	22 AAO09946	Human polypeptide
9	64.5	23.8	94	24 ABP75436	Human secretory po

10	64	23.6	339	22 AAU03138	Streptococcus pyog
11	64	23.6	339	23 ABP28098	Streptococcus poly
12	63.5	23.4	692	22 ABG09932	Novel human diagno
13	63.5	23.4	1798	22 ABG09933	Novel human diagno
14	63	23.2	60	22 AAO11745	Human polypeptide
15	62	22.9	228	23 ABP39849	Staphylococcus epi
16	62	22.9	228	23 ABP40650	Staphylococcus epi
17	62	22.9	340	24 ABU02808	S. pneumoniae type
18	62	22.9	619	23 AAO22174	Ramoplanin biosynt
19	61	22.5	510	22 ABB62105	Drosophila melanog
20	60	22.1	59	22 AAM82976	Human immune/haema
21	60	22.1	70	22 AAU18578	Human lung antigen
22	60	22.1	70	22 AAU17846	Novel human respir
23	60	22.1	3095	23 AAE20788	Rat C3b/C4b comple
24	59.5	22.0	463	22 AAB92562	Human protein sequ
25	59.5	22.0	725	20 AAY59724	Human normal ovari
26	59.5	22.0	877	22 AAB93668	Human protein sequ
27	59.5	22.0	1181	21 AAY82707	Human apoptosis re
28	59.5	22.0	1182	21 AAB42517	Human apoptosis re
29	59.5	22.0	1220	21 AAY82708	Human vesicle traf
30	59.5	22.0	1220	22 AAE04770	Novel human secret
31	59.5	22.0	1232	22 AAU32866	HEV-US1 ORF1 prote
32	59.5	22.0	1698	20 AAY31381	Novel human secret
33	59	21.8	86	22 AAU29913	Novel human secret
34	59	21.8	1012	21 AAY32241	Mouse tolloid-like
35	58.5	21.6	277	23 ABB47607	Listeria monocytog
36	58.5	21.6	458	22 ABG09851	Novel human diagno
37	58	21.4	63	22 AAM91439	Human immune/haema
38	58	21.4	160	22 AAU48056	Propionibacterium
39	57.5	21.2	576	23 ABP63134	lovF gene expressi
40	57.5	21.2	576	23 ABP35722	Fungal ZBC protein
41	57	21.0	55	21 AAB32094	Human secreted pro
42	57	21.0	84	22 AAO10311	Human polypeptide
43	57	21.0	102	22 AAM89598	Human immune/haema
44	57	21.0	1018	21 AAB07773	Large subunit of t
45	57	21.0	1018	22 AAB50543	B. lactofermentum

ALIGNMENTS

RESULT 1
AAB49223
ID AAB49223 standard; protein; 1500 AA.
XX AAB49223;
XX 13-MAR-2001 (first entry)
DT Human T1405 carbamyl phosphate synthase I DNA.
DE Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
XX sclerosis; pulmonary hypertension; bone marrow; gene therapy.
OS Homo sapiens.
XX WO2000073322-A1.
PN 07-DEC-2000.
PD 01-JUN-2000; 2000WO-US15079.
XX 01-JUN-1999; 99US-0323472.
PR (UIVA-) UNIV VANDERBILT.
XX Summar ML, Christman BW;
PI WPI; 2001-049926/06.
DR Detecting susceptibility to suboptimal urea cycle function, e.g. bone
XX marrow transplant toxicity, comprises identifying a polymorphism in the
PT gene for carbamylphosphate synthase

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 97735
; LENGTH: 2193
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-97735

Query Match 33.8%; Score 27.4; DB 13; Length 2193;
Best Local Similarity 59.7%; Pred. No. 2;
Matches 46; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 5 CATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCTCAAG 64
||||| ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 341 CATCTGAGTCTTTCACTCCGAATAAGCCCTCCCATCTTGAGACCTGCAAGGTGCAAG 282
| | | ||||| ||| |
QY 65 AAGGACAGAAATCCACG 81
| | | ||||| ||| |
Db 281 AGAGGTAGAAATCACATC 265

RESULT 14
US-10-027-632-103660/c
; Sequence 103660, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103660
; LENGTH: 2193
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-103660

Query Match 33.8%; Score 27.4; DB 13; Length 2193;
Best Local Similarity 59.7%; Pred. No. 2;
Matches 46; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 5 CATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCTCAAG 64
||||| ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 341 CATCTGAGTCTTTCACTCCGAATAAGCCCTCCCATCTTGAGACCTGCAAGGTGCAAG 282
| | | ||||| ||| |
QY 65 AAGGACAGAAATCCACG 81
| | | ||||| ||| |
Db 281 AGAGGTAGAAATCACATC 265

RESULT 15
US-10-207-655-78/c
; Sequence 78, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 78
; LENGTH: 6471
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-207-655-78

Query Match 33.8%; Score 27.4; DB 14; Length 6471;
Best Local Similarity 59.7%; Pred. No. 2.3;
Matches 46; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 5 CATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCTCAAG 64
||||| ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 332 CATCTGAGTCTTTCACTCCGAATAAGCCCTCCCATCTTGAGACCTGCAAGGTGCAAG 273
| | | ||||| ||| |
QY 65 AAGGACAGAAATCCACG 81
| | | ||||| ||| |
Db 272 AGAGGTAGAAATCACATC 256

Search completed: September 13, 2003, 02:11:07
Job time : 161 secs

Query Match	51.6%;	Score 41.8;	DB 10;	Length 242;
Best Local Similarity	95.6%;	Pred. NO. 7.8e-06;		
Matches 43;	Conservative	0;	Mismatches 2;	Indels 0;
	Gaps	0;		Gaps 0;

QY	37	GCCACCCAGTGGCATGCGCGTCTCAAGAAGGACAGAATCCCAGC	81
Db	1	GCTACCCAGTGGCATGCGCGTCTCAAGAAGGACAGAATCCCAGC	45

RESULT 10

```

US-09-801-274-1768
; Sequence 1768, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: 2825.2009-001
; CURRENT APPLICATION NUMBER: US/09/801,274
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1768
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-274-1768

```

Query Match	37.8%;	Score 30.6;	DB 9;	Length 31;
Best Local Similarity	96.8%;	Pred. No. 0.083;		
Matches 30; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

QY 26 ACAATGTCCTGCCACCCAGTGGCATGGCC 56
|||||:|||||
Db 1 ACAATGTCCTGCCAMCCAGTGGCATGGCC 31

RESULT 11

```

RESOLUTION II.
US-09-764-891-5507
; Sequence 5507, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5507
; LENGTH: 3591
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-5507

```

Query Match 36.3%; Score 29.4; DB 11; Length 3591;
Best Local Similarity 63.4%; Pred. No. 0.39;
Matches 45: Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY
10 GACTGGCTCAACGCCAACAAATGTCCCTGCCACCACCCAGTGGCATGGCGGCTCAAGAAGGA 69
||||| ||| | ||| ||||| | | ||||| ||| | ||| |
Db
1746 GACTGGCTGGACCCCACCGATGCTCCTGGCGTCACCCCCGCTCTGGCCCTCACTGGAGGCA 1805

QY 70 CAGAATCCAG 80
|||
Db 1806 GGGACACTCAG 1816

RESULT 12

US-10-068-674-l/c
; Sequence 1, Application US/10068674
; Publication No. US20020177203A1
; GENERAL INFORMATION:
; APPLICANT: Kevirikko, Kari
; APPLICANT: Pihlajaniemi, Taina A.
; TITLE OF INVENTION: 2 SUBUNIT OF PROLYL-4-HYDROXYLASE,
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE ENCODING SUCH SUBUNIT AND METHODS
; TITLE OF INVENTION: FOR PRODUCING THE SAME
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:

```

; ADDRESS: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
;

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/068,674
FILING DATE: 06-FEB-2002

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/633,879
FILING DATE: 10-MAR-1996
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 25,227
REFERENCE/DOCKET NUMBER: 8389-041
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-854-3660
TELEFAX: 415-854-3694
TELEX: 66141 PENNIE

```

/ IDENTITY: P00789-1
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2168 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: unknown
/ TOPOLOGY: unknown
/ MOLECULE TYPE: cDNA
/ FEATURE:

```

NAME/KEY: CDS
LOCATION: 151..1761
US-10-068-674-1

Query Match	34.6%;	Score 28;	DB 13;	Length 2168;
Best Local Similarity	63.2%;	Pred.No. 1.2;		
Matches 43:	Conservative	0;	Mismatches 25;	Indels 0;
	Gaps	0;		

QY	9	AGACTGGCTCAAGCCCAACAATGTCCCTGCCACCCCATGCGCATGGCGTCTCAAGAAGG	68
Dh	2005	AGTTGGCTCTACGCAGCCATGCCAAGCACCCCAGCGCCAGGCAGGTGTAGCTGG	1946

QY	69	ACAGAATC	76
Dh	1945	TCAGAGTC	1938

RESUME 13

```

US-10-027-632-97735/c
; Sequence 97735, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129

```

Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
|||||
Db 157 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 216
|||||

QY 61 CAAGAAGGACAGAAATCCCAGC 81
|||||
Db 217 CAAGAAGGACAGAAATCCCAGC 237
|||||

RESULT 6
US-10-101-510-553
; Sequence 553, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 553
; LENGTH: 5772
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (5240)..(5365)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-101-510-553

Query Match 100.0%; Score 81; DB 12; Length 5772;
Best Local Similarity 100.0%; Pred. No. 4e-20;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
|||||
Db 4310 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 4369
|||||

QY 61 CAAGAAGGACAGAAATCCCAGC 81
|||||
Db 4370 CAAGAAGGACAGAAATCCCAGC 4390
|||||

RESULT 7
US-09-880-107-1769
; Sequence 1769, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1769
; LENGTH: 5215
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D90282
US-09-880-107-1769

Query Match 98.0%; Score 79.4; DB 10; Length 5215;
Best Local Similarity 98.8%; Pred. No. 1.5e-19;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
|||||
Db 4295 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 4354
|||||

QY 61 CAAGAAGGACAGAAATCCCAGC 81
|||||
Db 4355 CAAGAAGGACAGAAATCCCAGC 4375
|||||

RESULT 8
US-09-960-352-12478
; Sequence 12478, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12478
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 53-LIB34-037-Q1-E1-F2
US-09-960-352-12478

Query Match 92.1%; Score 74.6; DB 10; Length 434;
Best Local Similarity 95.1%; Pred. No. 6.8e-18;
Matches 77; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
|||||
Db 36 GCCACATCAGACTGGCTCAATGCCAACAAATGTACCTGCTACCCAGTGGCATGGCCATCT 95
|||||

QY 61 CAAGAAGGACAGAAATCCCAGC 81
|||||
Db 96 CAAGAAGGACAGAAATCCCAGC 116
|||||

RESULT 9
US-09-960-352-1818
; Sequence 1818, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 1818
; LENGTH: 242
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 08-LIB34-052-Q1-E1-B11
US-09-960-352-1818


```
RESULT 2
US-09-902-941-93
; Sequence 93, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-93

Query Match      100.0%; Score 81; DB 10; Length 531;
Best Local Similarity 100.0%; Pred. No. 3e-20;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 216

QY 61 CAAGAAGGACAGATCCAGC 81
   ||||||||||||||||||
Db 217 CAAGAAGGACAGATCCAGC 237

RESULT 3
US-09-849-626-93
; Sequence 93, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-849-626-93

Query Match      100.0%; Score 81; DB 10; Length 531;
Best Local Similarity 100.0%; Pred. No. 3e-20;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 216

QY 61 CAAGAAGGACAGATCCAGC 81
   ||||||||||||||||||
Db 217 CAAGAAGGACAGATCCAGC 237

RESULT 4
US-10-017-754-93
; Sequence 93, Application US/09476300
; Publication No. US20030125245A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C3
; CURRENT APPLICATION NUMBER: US/09/476,300
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 785
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-017-754-93

Query Match      100.0%; Score 81; DB 11; Length 531;
Best Local Similarity 100.0%; Pred. No. 3e-20;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 216

QY 61 CAAGAAGGACAGATCCAGC 81
   ||||||||||||||||||
Db 217 CAAGAAGGACAGATCCAGC 237

RESULT 5
US-10-017-754-93
; Sequence 93, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-93

Query Match      100.0%; Score 81; DB 14; Length 531;
Best Local Similarity 100.0%; Pred. No. 3e-20;
```

```
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 216

QY 61 CAAGAAGGACAGATCCAGC 81
   ||||||||||||||||||
Db 217 CAAGAAGGACAGATCCAGC 237

RESULT 4
US-09-476-300-93
; Sequence 93, Application US/09476300
; Publication No. US20030125245A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C3
; CURRENT APPLICATION NUMBER: US/09/476,300
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 785
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-476-300-93

Query Match      100.0%; Score 81; DB 11; Length 531;
Best Local Similarity 100.0%; Pred. No. 3e-20;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 216

QY 61 CAAGAAGGACAGATCCAGC 81
   ||||||||||||||||||
Db 217 CAAGAAGGACAGATCCAGC 237

RESULT 5
US-10-017-754-93
; Sequence 93, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-93

Query Match      100.0%; Score 81; DB 14; Length 531;
Best Local Similarity 100.0%; Pred. No. 3e-20;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 13, 2003, 00:27:37 ; Search time 151 Seconds
(without alignments)
1302.306 Million cell updates/sec

Title: US-09-585-077C-3_COPY_4300_4380
Perfect score: 81
Sequence: 1 gccacatcagactggtctaa.....aagaaggacagaatcccagc 81

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1632420 seqs, 1213878141 residues

Total number of hits satisfying chosen parameters: 3264840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	531	10	US-09-736-457-93
2	81	100.0	531	10	US-09-902-941-93
3	81	100.0	531	10	US-09-849-626-93
4	81	100.0	531	11	US-09-476-300-93
5	81	100.0	531	14	US-10-017-754-93
6	81	100.0	5772	12	US-10-101-510-553
7	79.4	98.0	5215	10	US-09-880-107-1769
8	74.6	92.1	434	10	US-09-960-352-12478
9	41.8	51.6	242	10	US-09-960-352-1818
10	30.6	37.8	31	9	US-09-801-274-1768
11	29.4	36.3	3591	11	US-09-764-891-5507
12	28	34.6	2168	13	US-10-068-674-1
13	27.4	33.8	2193	13	US-10-027-632-97735
14	27.4	33.8	2193	13	US-10-027-632-103660
15	27.4	33.8	6471	14	US-10-207-655-78
16	27	33.3	288	10	US-09-983-965-4858

17	26.6	32.8	331	10	US-09-783-590-7357	Sequence 7357, Ap
c 18	26.4	32.6	969	10	US-09-280-197-13	Sequence 13, Appl
c 19	26.4	32.6	3276	10	US-09-280-197-4	Sequence 4, Appli
c 20	26.4	32.6	3276	10	US-09-423-126-8	Sequence 8, Appli
c 21	25.8	31.9	1281	14	US-10-156-761-2272	Sequence 2272, Ap
c 22	25.8	31.9	9025608	14	US-10-156-761-1	Sequence 1, Appli
c 23	25.6	31.6	469	11	US-09-918-995-2365	Sequence 2365, Ap
c 24	25.4	31.4	1117	13	US-10-027-632-116661	Sequence 116661,
c 25	25.4	31.4	1703	14	US-10-138-316-3	Sequence 3, Appli
c 26	25.4	31.4	113604	14	US-10-227-195A-1	Sequence 1, Appli
c 27	25.4	31.4	113604	14	US-10-227-195A-2	Sequence 2, Appli
c 28	25.2	31.1	567	11	US-09-918-995-26806	Sequence 26806, A
c 29	25	30.9	450	10	US-09-974-300-5526	Sequence 5526, Ap
c 30	24.8	30.6	284	10	US-09-833-381-417	Sequence 417, App
c 31	24.8	30.6	397	10	US-09-917-800A-897	Sequence 897, App
c 32	24.8	30.6	878	14	US-10-198-846-5229	Sequence 5229, Ap
c 33	24.8	30.6	957	13	US-10-027-632-163409	Sequence 163409,
c 34	24.8	30.6	18746	12	US-10-017-161-2229	Sequence 2229, Ap
c 35	24.6	30.4	723	13	US-10-027-632-165578	Sequence 165578,
c 36	24.4	30.1	375	14	US-10-106-698-3792	Sequence 3792, Ap
c 37	24.4	30.1	987	10	US-09-738-973-289	Sequence 289, App
c 38	24.4	30.1	987	10	US-09-854-133-289	Sequence 289, App
c 39	24.4	30.1	987	14	US-10-144-649A-289	Sequence 289, App
c 40	24.4	30.1	14917	11	US-09-909-567B-11	Sequence 11, Appli
c 41	24.4	30.1	146547	14	US-10-017-128-1	Sequence 1, Appli
c 42	24.2	29.9	14040	11	US-09-764-891-5478	Sequence 5478, Ap
c 43	24.2	29.9	14040	11	US-09-764-891-10205	Sequence 10205, A
c 44	24.2	29.9	14040	14	US-10-205-428-1004	Sequence 1004, Ap
c 45	24.2	29.9	28438	12	US-09-820-790-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-736-457-93
; Sequence 93, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Ligu
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-93

Query Match	100.0%;	Score 81;	DB 10;	Length 531;
Best Local Similarity	100.0%;	Pred. No. 3e-20;		
Matches	81;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	GCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCT	60	
Db	157	GCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCT	216	
QY	61	CAAGAAGGACAGATCCCAGC	81	
Db	217	CAAGAAGGACAGATCCCAGC	237	

```
; EARLIER FILING DATE: 1998-08-17
; EARLIER APPLICATION NUMBER: 60/094,477
; EARLIER FILING DATE: 1998-07-29
; EARLIER APPLICATION NUMBER: 08/921,068
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 08/739,383
; EARLIER FILING DATE: 1996-10-29
; EARLIER APPLICATION NUMBER: 60/019,014
; EARLIER FILING DATE: 1995-12-22
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 3
; LENGTH: 1703
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (193)..(579)
US-09-597-735-3

Query Match      31.4%; Score 25.4; DB 4; Length 1703;
Best Local Similarity 68.6%; Pred. No. 5.2;
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY      3 CACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATG 53
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1631 CTCACAAAACCGGCCCATCACCTCAAAAGGACCTGCCACCCAGTGTATG 1581

Search completed: September 13, 2003, 00:29:29
Job time : 51.5 secs
```

```
; APPLICANT: Christman, Brian
; TITLE OF INVENTION: HUMAN CARBAMYL PHOSPHATE SYNTHETASE I POLYMORPHISM AND DIAGNOST
; TITLE OF INVENTION: THERAPEUTIC METHODS RELATING THERETO
; FILE REFERENCE: 1242/19
; CURRENT APPLICATION NUMBER: US/09/323,472A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 5761
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)..(4626)
US-09-323-472A-11

Query Match          98.0%; Score 79.4; DB 4; Length 5761;
Best Local Similarity 98.8%; Pred. No. 1.8e-19;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCGGTCT 60
   |||||
Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCGGTCT 4359

QY 61 CAAGAAGGACAGATCCCGC 81
   |||||
Db 4360 CAAGAAGGACAGATCCCGC 4380

RESULT 7
US-09-323-472A-13
; Sequence 13, Application US/09323472A
; Patent No. 6346382
; GENERAL INFORMATION:
; APPLICANT: Summar, Marshall
; APPLICANT: Christman, Brian
; TITLE OF INVENTION: HUMAN CARBAMYL PHOSPHATE SYNTHETASE I POLYMORPHISM AND DIAGNOST
; TITLE OF INVENTION: THERAPEUTIC METHODS RELATING THERETO
; FILE REFERENCE: 1242/19
; CURRENT APPLICATION NUMBER: US/09/323,472A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 5762
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)..(4626)
US-09-323-472A-13

Query Match          98.0%; Score 79.4; DB 4; Length 5762;
Best Local Similarity 98.8%; Pred. No. 1.8e-19;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCGGTCT 60
   |||||
Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCGGTCT 4359

QY 61 CAAGAAGGACAGATCCCGC 81
   |||||
Db 4360 CAAGAAGGACAGATCCCGC 4380

RESULT 8
US-08-633-879C-1/c
; Sequence 1, Application US/08633879C
; Patent No. 5928922
; GENERAL INFORMATION:
; APPLICANT: Kivirikko, Kari I.
; APPLICANT: Pihlajaniemi, Taina
```

```
; APPLICANT: Helaakoski, Tarja I.
; APPLICANT: Annunen, Pia P.
; APPLICANT: Nissi, Ritva K.
; APPLICANT: No. 5928922elainen, Minna K.
; TITLE OF INVENTION: '2 SUBUNIT OF PROLYL-4-HYDROXYLASE
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING SUCH SUBUNIT AND
; TITLE OF INVENTION: METHODS FOR PRODUCING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/633,879C
; FILING DATE: 10-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8389-0041-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2168 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 151...1761
; OTHER INFORMATION:
US-08-633-879C-1

Query Match          34.6%; Score 28; DB 2; Length 2168;
Best Local Similarity 53.2%; Pred. No. 0.63;
Matches 43; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 9 AGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCGGTCTCAAGAAG 68
   |||||
Db 2005 AGTCTGGCTCTACCGCAGCCATGCCAAGACACCCAGAGCCAGGCGGTAGCTGG 1946
   |||||
QY 69 ACAGAAATC 76
   |||||
Db 1945 TCAGAGTC 1938

RESULT 9
US-08-633-768A-4/c
; Sequence 4, Application US/08633768A
; Patent No. 6013504
; GENERAL INFORMATION:
; APPLICANT: YU, SHUKUN
; APPLICANT: BOJSEN, KIRSTEN
; APPLICANT: KRAGH, KARSTEN
; APPLICANT: BOJKO, MAJA
; APPLICANT: NIELSEN, JOHN
; APPLICANT: MARCUSSEN, JAN
; TITLE OF INVENTION: ALPHA-1,4-GLUCAN LYASE FROM
; TITLE OF INVENTION: A FUNGUS INFECTED ALGAE, ITS PURIFICATION
```



```
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-93

Query Match          100.0%; Score 81; DB 4; Length 531;
Best Local Similarity 100.0%; Pred. No. 2.3e-20;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 216

QY      61 CAAGAAGGACAGAAATCCAGC 81
      ||||||||||||||||||
Db      217 CAAGAAGGACAGAAATCCAGC 237

RESULT 3
US-09-736-457-93
; Sequence 93, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-93

Query Match          100.0%; Score 81; DB 4; Length 531;
Best Local Similarity 100.0%; Pred. No. 2.3e-20;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 216

QY      61 CAAGAAGGACAGAAATCCAGC 81
      ||||||||||||||||||
Db      217 CAAGAAGGACAGAAATCCAGC 237

RESULT 4
US-09-323-472A-3
; Sequence 3, Application US/09323472A
; Patent No. 6346382
; GENERAL INFORMATION:
; APPLICANT: Summar, Marshall
```

```
; APPLICANT: Summar, Marshall
; APPLICANT: Christman, Brian
; TITLE OF INVENTION: HUMAN CARBAMYL PHOSPHATE SYNTHETASE I POLYMORPHISM AND DIAGN
; TITLE OF INVENTION: THERAPEUTIC METHODS RELATING THERETO
; FILE REFERENCE: 1242/19
; CURRENT APPLICATION NUMBER: US/09/323,472A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 5761
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)..(4626)
US-09-323-472A-3

Query Match          100.0%; Score 81; DB 4; Length 5761;
Best Local Similarity 100.0%; Pred. No. 4.8e-20;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 4359

QY      61 CAAGAAGGACAGAAATCCAGC 81
      ||||||||||||||||||
Db      4360 CAAGAAGGACAGAAATCCAGC 4380

RESULT 5
US-09-323-472A-1
; Sequence 1, Application US/09323472A
; Patent No. 6346382
; GENERAL INFORMATION:
; APPLICANT: Summar, Marshall
; APPLICANT: Christman, Brian
; TITLE OF INVENTION: HUMAN CARBAMYL PHOSPHATE SYNTHETASE I POLYMORPHISM AND DIAGN
; TITLE OF INVENTION: THERAPEUTIC METHODS RELATING THERETO
; FILE REFERENCE: 1242/19
; CURRENT APPLICATION NUMBER: US/09/323,472A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 5761
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)..(4626)
US-09-323-472A-1

Query Match          98.0%; Score 79.4; DB 4; Length 5761;
Best Local Similarity 98.8%; Pred. No. 1.8e-19;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 4359

QY      61 CAAGAAGGACAGAAATCCAGC 81
      ||||||||||||||||||
Db      4360 CAAGAAGGACAGAAATCCAGC 4380

RESULT 6
US-09-323-472A-11
; Sequence 11, Application US/09323472A
; Patent No. 6346382
; GENERAL INFORMATION:
; APPLICANT: Summar, Marshall
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 22:27:31 ; Search time 44.5 Seconds
(without alignments)
803.416 Million cell updates/sec

Title: US-09-585-077C-3_COPY_4300_4380
Perfect score: 81
Sequence: 1 gccacatcagactggtctcaa.....aagaaggacagaatccagc 81

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/pctus_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	495	4 US-09-323-472A-5	Sequence 5, Appli
2	81	100.0	531	4 US-09-702-705-93	Sequence 93, Appl
3	81	100.0	531	4 US-09-736-457-93	Sequence 93, Appl
4	81	100.0	5761	4 US-09-323-472A-3	Sequence 3, Appli
5	79.4	98.0	5761	4 US-09-323-472A-1	Sequence 1, Appli
6	79.4	98.0	5761	4 US-09-323-472A-11	Sequence 11, Appl
7	79.4	98.0	5762	4 US-09-323-472A-13	Sequence 13, Appl
8	28	34.6	2168	2 US-08-633-879C-1	Sequence 1, Appli
9	26.4	32.6	3276	3 US-08-633-768A-4	Sequence 4, Appli
10	25.6	31.6	468	3 US-09-067-782A-6	Sequence 6, Appli
11	25.4	31.4	1703	3 US-09-135-021-77	Sequence 77, Appl
12	25.4	31.4	1703	3 US-09-135-020-3	Sequence 3, Appli
13	25.4	31.4	1703	3 US-09-135-010A-3	Sequence 3, Appli
14	25.4	31.4	1703	4 US-09-444-871-3	Sequence 3, Appli
15	25.4	31.4	1703	4 US-09-597-735-3	Sequence 3, Appli
16	25.4	31.4	1703	4 US-09-444-295-3	Sequence 3, Appli
17	25.4	31.4	1703	4 US-09-597-732-3	Sequence 3, Appli
18	25.4	31.4	1703	4 US-09-597-731-3	Sequence 3, Appli
19	24.8	30.6	654	4 US-09-252-991A-12606	Sequence 12606, A
20	24.6	30.4	2265	2 US-08-940-332-1	Sequence 1, Appli
21	24.6	30.4	4403765	3 US-09-103-840A-2	Sequence 2, Appli
22	24.6	30.4	4411529	3 US-09-103-840A-1	Sequence 1, Appli
23	24.4	30.1	987	4 US-09-370-838-289	Sequence 289, App
24	24	29.6	432	4 US-09-252-991A-16133	Sequence 16133, A
25	24	29.6	438	4 US-09-252-991A-8292	Sequence 8292, Ap
26	24	29.6	1036	4 US-09-016-434-1308	Sequence 1308, Ap
27	23.6	29.1	1056	3 US-09-067-782A-1	Sequence 1, Appli

c 28	23.6	29.1	1520	5 PCT-US93-07213-8	Sequence 8, Appli
c 29	23.6	29.1	2219	5 PCT-US93-07213-4	Sequence 4, Appli
c 30	23.6	29.1	2870	5 PCT-US93-07213-3	Sequence 3, Appli
c 31	23.6	29.1	3102	5 PCT-US93-07213-1	Sequence 1, Appli
c 32	23.6	29.1	152331	3 US-09-128-155-16	Sequence 16, Appl
c 33	23.4	28.9	519	4 US-09-844-634-19	Sequence 19, Appl
c 34	23.4	28.9	1886	4 US-09-149-476-93	Sequence 93, Appl
c 35	23.4	28.9	3683	4 US-09-844-634-3	Sequence 3, Appli
c 36	23.4	28.9	15602	4 US-09-844-634-17	Sequence 17, Appl
c 37	23.2	28.6	1900	4 US-09-608-285A-47	Sequence 47, Appl
c 38	23.2	28.6	1900	4 US-09-557-800C-47	Sequence 47, Appl
c 39	23.2	28.6	2294	4 US-09-608-285A-49	Sequence 49, Appl
c 40	23.2	28.6	2294	4 US-09-557-800C-49	Sequence 49, Appl
c 41	23.2	28.6	2371	4 US-09-608-285A-46	Sequence 46, Appl
c 42	23.2	28.6	2371	4 US-09-557-800C-46	Sequence 46, Appl
c 43	23.2	28.6	2497	4 US-09-608-285A-51	Sequence 51, Appl
c 44	23.2	28.6	2497	4 US-09-557-800C-51	Sequence 51, Appl
c 45	23.2	28.6	2693	4 US-09-608-285A-48	Sequence 48, Appl

ALIGNMENTS

RESULT 1

US-09-323-472A-5
; Sequence 5, Application US/09323472A
; Patent No. 6346382
; GENERAL INFORMATION:
; APPLICANT: Summar, Marshall
; APPLICANT: Christman, Brian
; TITLE OF INVENTION: HUMAN CARBAMYL PHOSPHATE SYNTHETASE I POLYMORPHISM AND DIAGN
; TITLE OF INVENTION: THERAPEUTIC METHODS RELATING THERETO
; FILE REFERENCE: 1242/19
; CURRENT APPLICATION NUMBER: US/09/323,472A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (111)..(224)
; OTHER INFORMATION: n is G or A or C or T/U
US-09-323-472A-5

Query Match	100.0%;	Score 81;	DB 4;	Length 495;
Best Local Similarity	100.0%;	Pred. No. 2.3e-20;		
Matches	81;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCT 60		
Db	126	GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCT 185		
Qy	61	CAAGAAGGACAGATCCCAGC 81		
Db	186	CAAGAAGGACAGATCCCAGC 206		

RESULT 2

US-09-702-705-93
; Sequence 93, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane

DEFINITION 602718213F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4858442 5',
mRNA sequence.
ACCESSION BG761337
VERSION BG761337.1 GI:14071990
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 699)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1712 row: i column: 03
High quality sequence stop: 695.
Location/Qualifiers
FEATURES
source
1..699
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4858442"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_49"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 185 a 166 c 158 g 190 t
ORIGIN
Query Match 98.0%; Score 79.4; DB 12; Length 699;
Best Local Similarity 98.8%; Pred. No. 2.3e-14;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCCGCATGGCCGTCT 60
|||||
Db 517 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCCGCATGGCCGTCT 576
QY 61 CAAGAAGGACAGAAATCCACG 81
|||||
Db 577 CAAGAAGGACAGAAATCCACG 597
RESULT 15
BG616938
LOCUS 602615612F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4734387 5',
mRNA sequence.
DEFINITION BG616938
ACCESSION BG616938
VERSION BG616938.1 GI:13668309
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 706)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1597 row: h column: 04
High quality sequence stop: 706.
Location/Qualifiers
FEATURES
source
1..706
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4734387"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_76"
/note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgcctcgcc); Site_2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCCCATATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGCGGCGCATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.85
kb (range 1.0-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
BASE COUNT 200 a 173 c 151 g 181 t 1 others
ORIGIN
Query Match 98.0%; Score 79.4; DB 10; Length 706;
Best Local Similarity 98.8%; Pred. No. 2.3e-14;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCCGCATGGCCGTCT 60
|||||
Db 283 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCCGCATGGCCGTCT 342
QY 61 CAAGAAGGACAGAAATCCACG 81
|||||
Db 343 CAAGAAGGACAGAAATCCACG 363
Search completed: September 13, 2003, 00:27:27
Job time : 1961 secs

JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1597 row: h column: 04
High quality sequence stop: 706.
Location/Qualifiers
FEATURES
source
1..706
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4734387"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_76"
/note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgcctcgcc); Site_2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCCCATATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGCGGCGCATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.85
kb (range 1.0-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
BASE COUNT 200 a 173 c 151 g 181 t 1 others
ORIGIN
Query Match 98.0%; Score 79.4; DB 10; Length 706;
Best Local Similarity 98.8%; Pred. No. 2.3e-14;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCCGCATGGCCGTCT 60
|||||
Db 283 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCCGCATGGCCGTCT 342
QY 61 CAAGAAGGACAGAAATCCACG 81
|||||
Db 343 CAAGAAGGACAGAAATCCACG 363
Search completed: September 13, 2003, 00:27:27
Job time : 1961 secs

```
source
1. .362
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCERB10"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT      106 a      92 c      71 g      93 t
ORIGIN

Query Match      98.8%; Score 80; DB 9; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.2e-14;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 CCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCTC 61
        |||||||
Db      1 CCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCTC 60

QY      62 AAGAAGGACAGATCCCCAGC 81
        |||||||
Db      61 AAGAAGGACAGATCCCCAGC 80

RESULT 12
AV660978      365 bp      mRNA      linear      EST 16-JAN-2002
LOCUS      AV660978      GLC Homo sapiens cDNA clone GLGND01 3', mRNA sequence.
DEFINITION      AV660978
ACCESSION      AV660978
VERSION      AV660978.1 GI:9881992
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 365)
AUTHORS      Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
JOURNAL      21625106
MEDLINE      11752456
PUBMED
COMMENT      Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1. .365
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLGND01"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT      90 a      97 c      86 g      92 t
ORIGIN

Query Match      98.0%; Score 79.4; DB 9; Length 365;
```

```
Best Local Similarity 98.8%; Pred. No. 1.8e-14;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
        |||||||
Db      271 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGTCAACCCAGTGGCATGGCCGTCT 330

QY      61 CAAGAAGGACAGAAATCCCAGC 81
        |||||||
Db      331 CAAGAAGGACAGAAATCCCAGC 351

RESULT 13
AV653332      626 bp      mRNA      linear      EST 15-JAN-2002
LOCUS      AV653332      GLC Homo sapiens cDNA clone GLCDJG10 3', mRNA sequence.
DEFINITION      AV653332
ACCESSION      AV653332
VERSION      AV653332.1 GI:9874346
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 626)
AUTHORS      Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
JOURNAL      21625106
MEDLINE      11752456
PUBMED
COMMENT      Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1. .626
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCDJG10"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT      156 a      152 c      142 g      175 t      1 others
ORIGIN

Query Match      98.0%; Score 79.4; DB 9; Length 626;
Best Local Similarity 98.8%; Pred. No. 2.2e-14;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
        |||||||
Db      468 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 527

QY      61 CAAGAAGGACAGAAATCCCAGC 81
        |||||||
Db      528 CAAGAAGGACAGAAATCCCAGC 548

RESULT 14
BG761337      699 bp      mRNA      linear      EST 15-MAY-2001
LOCUS      BG761337
```


LOCUS AL046242 646 bp mRNA linear EST 29-FEB-2000
DEFINITION DKFZp434E087_s1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION DKFZp434E087 3', mRNA sequence.
AL046242
VERSION AL046242.1 GI:5434326
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 646)
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE EST (Koehrer, et al.)
JOURNAL Unpublished
COMMENT Contact: Koehrer K
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
r1 sequence also available.
This clone (DKFZp434E087) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
Location/Qualifiers
1..646
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434E087"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 171 a 128 c 156 g 190 t 1 others
ORIGIN
|||||
1..646
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434E087"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 171 a 128 c 156 g 190 t 1 others
ORIGIN
|||||

Query Match 100.0%; Score 81; DB 9; Length 646;
Best Local Similarity 100.0%; Pred. No. 7e-15;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAAGCCCAACAATGTCCCTGCCACCCAGTGGCGCGTCT 60
|||||
Db 469 GCCACATCAGACTGGCTCAAGCCCAACAATGTCCCTGCCACCCAGTGGCGCGTCT 410
|||||

QY 61 CAAGAAGGACAGATCCAGC 81
|||||
Db 409 CAAGAAGGACAGATCCAGC 389
|||||

RESULT 10
BG533520
LOCUS 949 bp mRNA linear EST 03-APR-2001
DEFINITION BG533520 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4071944 5',
mRNA sequence.
ACCESSION BG533520
VERSION BG533520.1 GI:13525060
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 949)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.

LOCUS AV656423 362 bp mRNA linear EST 16-JAN-2002
DEFINITION AV656423 GLC Homo sapiens cDNA clone GLCERB10 3', mRNA sequence.
ACCESSION AV656423
VERSION AV656423.1 GI:9877437
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 362)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
JOURNAL 21625106
MEDLINE 11752456
PUBMED
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Location/Qualifiers

cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM918 row: f column: 09
High quality sequence stop: 740.

FEATURES
source
Location/Qualifiers
1..949
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4071944"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_76"
/note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgctcgcc); Site_2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCCATATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGGCACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.85
kb (range 1.0-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

BASE COUNT 271 a 224 c 181 g 272 t 1 others
ORIGIN
|||||

Query Match 100.0%; Score 81; DB 10; Length 949;
Best Local Similarity 100.0%; Pred. No. 8e-15;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAAGCCCAACAATGTCCCTGCCACCCAGTGGCGCGTCT 60
|||||
Db 181 GCCACATCAGACTGGCTCAAGCCCAACAATGTCCCTGCCACCCAGTGGCGCGTCT 240
|||||

QY 61 CAAGAAGGACAGATCCAGC 81
|||||
Db 241 CAAGAAGGACAGATCCAGC 261
|||||

RESULT 11
AV656423
LOCUS 362 bp mRNA linear EST 16-JAN-2002
DEFINITION AV656423 GLC Homo sapiens cDNA clone GLCERB10 3', mRNA sequence.
ACCESSION AV656423
VERSION AV656423.1 GI:9877437
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 362)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
JOURNAL 21625106
MEDLINE 11752456
PUBMED
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Location/Qualifiers

/mol_type="mrna"
/db_xref="taxon:9606"
/clone="LI7N670205-7-B10"
/sex="F"
/lab_host="Top10F"
/clone_lib="LI7N670205"
/note="Organ: Liver; Vector: pT7T3-Pac; Site_1: EcoRI;
Site_2: NotI; The library was contributed by the Soares
laboratory and it was constructed as described by Bonaldo,
M.P., Lennon, G. and Soares, M.B. (1996), Genome Research
6(9): 791-806. RNA was prepared from harvested cell
culture."

BASE COUNT 178 a 159 c 134 g 154 t
ORIGIN

Query Match 100.0%; Score 81; DB 14; Length 625;
Best Local Similarity 100.0%; Pred. No. 7e-15;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAAGCCCAACAATGTCCTGCCACCCAGTGGCGGTCT 60
|||||
Db 221 GCCACATCAGACTGGCTCAAGCCCAACAATGTCCTGCCACCCAGTGGCGGTCT 280
|||||
QY 61 CAAGAAGGACAGAATCCAGC 81
|||||
Db 281 CAAGAAGGACAGAATCCAGC 301
|||||

RESULT 7
AV654671 636 bp mRNA linear EST 15-JAN-2002
LOCUS
DEFINITION AV654671 GLC Homo sapiens cDNA clone GLCDYC09 3', mRNA sequence.
ACCESSION AV654671
VERSION AV654671.1 GI:9875685
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 636)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.

Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
PUBMED 11752456
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

Location/Qualifiers

1. .636
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="GLCDYC09"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

FEATURES
source

Query Match 100.0%; Score 81; DB 14; Length 625;
Best Local Similarity 100.0%; Pred. No. 7e-15;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

BASE COUNT 179 a 148 c 122 g 185 t 2 others
ORIGIN

Query Match 100.0%; Score 81; DB 9; Length 636;
Best Local Similarity 100.0%; Pred. No. 7e-15;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAAGCCCAACAATGTCCTGCCACCCAGTGGCGGTCT 60
|||||
Db 59 GCCACATCAGACTGGCTCAAGCCCAACAATGTCCTGCCACCCAGTGGCGGTCT 118
|||||
QY 61 CAAGAAGGACAGAATCCAGC 81
|||||
Db 119 CAAGAAGGACAGAATCCAGC 139
|||||

RESULT 8
AV683932 637 bp mRNA linear EST 16-JAN-2002
LOCUS
DEFINITION AV683932 GKC Homo sapiens cDNA clone GKCDID08 5', mRNA sequence.
ACCESSION AV683932
VERSION AV683932.1 GI:10285795
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 637)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.

Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
PUBMED 11752456
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

Location/Qualifiers

1. .637
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="GKCDID08"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GKC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 173 a 147 c 118 g 196 t 3 others
ORIGIN

Query Match 100.0%; Score 81; DB 9; Length 637;
Best Local Similarity 100.0%; Pred. No. 7e-15;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAAGCCCAACAATGTCCTGCCACCCAGTGGCGGTCT 60
|||||
Db 5 GCCACATCAGACTGGCTCAAGCCCAACAATGTCCTGCCACCCAGTGGCGGTCT 64
|||||
QY 61 CAAGAAGGACAGAATCCAGC 81
|||||
Db 65 CAAGAAGGACAGAATCCAGC 85
|||||

RESULT 9
AL046242/c

REFERENCE 1 (bases 1 to 540)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 14 row: F column: 07
High quality sequence stop: 540.
FEATURES
source
1. .540
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S20T665307-14-F07"
/sex="M"
/lab_host="Top10F"
/clone_lib="S20T665307"
/note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."
BASE COUNT 147 a 133 c 121 g 139 t
ORIGIN
Query Match 100.0%; Score 81; DB 12; Length 540;
Best Local Similarity 100.0%; Pred. No. 6.6e-15;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCT 60
|||||
Db 329 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCT 388
QY 61 CAAGAAGGACAGAAATCCCAGC 81
|||||
Db 389 CAAGAAGGACAGAAATCCCAGC 409
RESULT 5
AV661502
LOCUS AV661502 GLC Homo sapiens cDNA clone GLCGTB04 3', mRNA sequence. EST 16-JAN-2002
DEFINITION AV661502
ACCESSION AV661502
VERSION AV661502.1 GI:9882516
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 564)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,

TITLE Hu,G., Gu,J., Chen,Z. and Han,Z.
Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE 21625106
PUBMED 11752456
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1. .564
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCGTB04"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 140 a 140 c 129 g 155 t
ORIGIN
Query Match 100.0%; Score 81; DB 9; Length 564;
Best Local Similarity 100.0%; Pred. No. 6.7e-15;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCT 60
|||||
Db 453 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCT 512
QY 61 CAAGAAGGACAGAAATCCCAGC 81
|||||
Db 513 CAAGAAGGACAGAAATCCCAGC 533
RESULT 6
CB153827
LOCUS CB153827 625 bp mRNA linear EST 29-JAN-2003
DEFINITION X-EST0211442 L17N670205 Homo sapiens cDNA clone L17N670205-7-B10 5', mRNA sequence.
ACCESSION CB153827
VERSION CB153827.1 GI:28138823
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 625)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 7 row: B column: 10
High quality sequence stop: 625.
FEATURES
source
1. .625
/organism="Homo sapiens"

Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source

Location/Qualifiers
1..366
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GKCBUC07"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GKC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 92 a 98 c 82 g 93 t 1 others
ORIGIN

Query Match 100.0%; Score 81; DB 9; Length 366;
Best Local Similarity 100.0%; Pred. No. 5.8e-15;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCT 60
193 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCT 252
Db
QY 61 CAAGAAGGACAGATCCCAGC 81
253 CAAGAAGGACAGATCCCAGC 273
Db

RESULT 2
AV661012 374 bp mRNA linear EST 16-JAN-2002
LOCUS AV661012 GLC Homo sapiens cDNA clone GLCGNF11 3', mRNA sequence.
DEFINITION AV661012
ACCESSION AV661012
VERSION AV661012.1 GI:9882026
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 374)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE 21625106
PUBMED 11752456
COMMENT Contact: Zeguap Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source

Location/Qualifiers
1..374
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCGNF11"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 91 a 103 c 87 g 93 t
ORIGIN
Query Match 100.0%; Score 81; DB 9; Length 374;
Best Local Similarity 100.0%; Pred. No. 5.8e-15;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCT 60
277 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCT 336
Db
QY 61 CAAGAAGGACAGATCCCAGC 81
337 CAAGAAGGACAGATCCCAGC 357
Db

RESULT 3
AII74790 501 bp mRNA linear EST 11-NOV-1999
LOCUS AII74790
DEFINITION HA2511 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
sequence.
ACCESSION AII74790
VERSION AII74790.1 GI:6361174
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 501)
AUTHORS Yu,X., Zhang,C., Luo,L., Ouyang,S., Li,W., Wu,J., Zhou,S., Liu,M.
and He,F.
TITLE Expression profile analysis of a human fetal liver cDNA library
JOURNAL Unpublished
COMMENT Contact: Chenggang Zhang
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Email: zhang.chenggang@hotmail.com.
FEATURES Location/Qualifiers
1..501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="MCL061/P3"
/clone_lib="Human fetal liver cDNA library"
/note="Vector: pCDNA1"

BASE COUNT 143 a 124 c 110 g 124 t
ORIGIN
Query Match 100.0%; Score 81; DB 9; Length 501;
Best Local Similarity 100.0%; Pred. No. 6.4e-15;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCT 60
181 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCGTCT 240
Db
QY 61 CAAGAAGGACAGATCCCAGC 81
241 CAAGAAGGACAGATCCCAGC 261
Db

RESULT 4
BM821786 540 bp mRNA linear EST 06-MAR-2002
LOCUS BM821786
DEFINITION K-EST0091028 S20T665307 Homo sapiens cDNA clone S20T665307-14-F07
5', mRNA sequence.
ACCESSION BM821786
VERSION BM821786.1 GI:19178199
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 22:26:16 ; Search time 1960 seconds
(without alignments)
1004.420 Million cell updates/sec

Title: US-09-585-077C-3_COPY_4300_4380
Perfect score: 81
Sequence: 1 gccacatcagactgctcaa.....aagaaggacagaatcccagc 81

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	81	100.0 366 9	AV696724 AV696724
2	81	100.0 374 9	AV661012 AV661012
3	81	100.0 501 9	AI174790 HA2511 Hu
4	81	100.0 540 12	BM821786 K-EST0091

5	81	100.0	564	9	AV661502	AV661502
6	81	100.0	625	14	CB153827	CB153827 K-EST0211
7	81	100.0	636	9	AV654671	AV654671
8	81	100.0	637	9	AV683932	AV683932
9	81	100.0	646	9	AL046242	DKFZp434E
10	81	100.0	949	10	BG533520	BG533520 601860935
11	80	98.8	362	9	AV656423	AV656423
12	79.4	98.0	365	9	AV660978	AV660978
13	79.4	98.0	626	9	AV653332	AV653332
14	79.4	98.0	699	12	BG761337	BG761337 602718213
15	79.4	98.0	706	10	BG616938	BG616938 602615612
16	79.4	98.0	843	10	BE971350	BE971350 601651514
17	78.4	96.8	490	9	AI065054	AI065054 HA0887 Hu
18	76.2	94.1	444	9	AA883205	AA883205 am16h02.s
19	76.2	94.1	830	10	BG569284	BG569284 602588576
20	74.8	92.3	405	14	T59186	T59186 yb50e08.r1
21	74.6	92.1	350	14	CB781273	CB781273 AMGNNUC:C
22	74.6	92.1	554	12	BM432173	BM432173 LJEJ1C10.
23	74.6	92.1	688	14	CB462893	CB462893 723232 MA
24	71.4	88.1	473	9	AI786068	AI786068 uJ57h08.Y
25	71.4	88.1	492	9	AA250015	AA250015 mz59d09.r
26	71.4	88.1	524	9	AA268939	AA268939 vb01g03.r
27	71.4	88.1	629	9	AA986202	uc73e03.Y
28	71.4	88.1	706	14	CB950475	CB950475 AGENCOURT
29	71.4	88.1	740	14	CB948817	CB948817 AGENCOURT
30	71.4	88.1	791	12	BI217440	BI217440 602933642
31	71.4	88.1	920	10	BF533080	BF533080 602073616
32	71.4	88.1	1057	14	W29382	W29382 mb99c06.r1
33	71.4	88.1	3161	11	AK028683	AK028683 Mus muscu
34	70.6	87.2	307	14	T77346	T77346 yd72e05.r1
35	70.4	86.9	725	12	BI247566	BI247566 602960022
36	69.8	86.2	511	9	AA511517	AA511517 vJ28c06.r
37	69	85.2	541	9	AA393420	AA393420 zt75f02.r
38	69	85.2	553	9	AA398737	AA398737 zt75f02.s
39	68.6	84.7	512	12	BQ011349	BQ011349 UI-1-BC1p
40	67.4	83.2	521	9	AA190971	AA190971 zp83h12.r
41	67	82.7	391	9	AV655328	AV655328
42	65.6	81.0	622	9	AI047399	AI047399 ud65g10.Y
43	62	76.5	564	14	CB120338	CB120338 K-EST0167
44	61.6	76.0	372	10	BF547163	BF547163 UI-R-Cl-1
45	60	74.1	745	14	CB594152	CB594152 AGENCOURT

ALIGNMENTS

RESULT 1
LOCUS AV696724 366 bp mRNA linear EST 16-JAN-2002
DEFINITION AV696724 GK Homo sapiens cDNA clone GKCBUC07 5', mRNA sequence.
ACCESSION AV696724
VERSION AV696724.1 GI:10298587
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 366)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE 21625106
PUBMED 11752456
COMMENT Contact: Zeguag Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)

QY 61 CACGAAGGACAGATCCCGC 81
| | | | | | | | | | | | | | | |
Db 96 CACGAAGGACAGATCCCGC 116

Search completed: September 12, 2003, 22:34:27
Job time : 203.5 secs

XX Toxicologically relevant human nucleotide sequence #1900.
DE
XX Toxicologically relevant gene; toxicological response; gene; ss.
KW Homo sapiens.
XX WO2003016500-A2.
PN
XX
PD 27-FEB-2003.
XX
PF 16-AUG-2002; 2002WO-US26514.
XX
PR 16-AUG-2001; 2001US-313080P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Neft RE, Dunn RT, Adkins K, Pickett GG, Kier LD, Schmeiser K;
PI Alen P;
XX
DR WPI; 2003-268322/26.
XX
PT Determining a toxicological response to an agent, useful for screening
PT of drugs, comprises comparing the expression profile of one or more
PT human toxic response genes to a reference gene expression profile
PT indicative of toxicity -
XX
PS Claim 1; Page 447; 455pp; English.
XX
CC The present invention describes a method (M1) for determining a
CC toxicological response to an agent, which comprises comparing the
CC expression profile of one or more human toxic response genes to a
CC reference gene expression profile indicative of toxicity, and so
CC determining the presence of a toxic response to the agent. Also
CC described: (1) an array comprising one or more polynucleotides selected
CC from the genes corresponding to the partial sequences given in ABZ82842
CC to ABZ84764, or their fragments of at least 20 nucleotides, or
CC homologues; and (2) determining if a gene putatively identified to be a
CC toxic response gene plays a role on toxic response pathways by
CC determining the expression profile of the gene after exposure of cells
CC or a human subject to a known toxic pharmaceutical or industrial agent,
CC comprising: (a) exposing cells to an agent or isolating cells from a
CC human subject who was exposed to an agent; (b) obtaining the test gene
CC expression profile for a putatively identified toxic response gene after
CC exposure to a known toxic pharmaceutical or industrial agent; and
CC (c) comparing the test profile to the expression profile of a gene with
CC a similar function or comparing the test profile to the expression
CC profile of that gene after exposure to other known toxic compounds. The
CC methods are useful for predicting and determining toxicological responses
CC on a cellular, organ or system level. The arrays comprising the human
CC genes are useful for toxicological screening of drugs, pharmaceutical
CC compounds and chemicals.
XX
SQ Sequence 715 BP; 196 A; 173 C; 141 G; 205 T; 0 other;

Query Match 94.8%; Score 76.8; DB 25; Length 715;
Best Local Similarity 97.5%; Pred. No. 4.5e-17;
Matches 78; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCACATCAGACTGGCTCAAGCCCAACAATGTCCCTGCCACCCAGTGGCATGGCGTCTC 61
|| |||||
Db 49 CCTATCAGACTGGCTCAAGCCCAACAATGTCCCTGCCACCCAGTGGCATGGCGTCTC 108

QY 62 AGAAGGACAGATCCACG 81
|||||
Db 109 AGAAGGACAGATCCACG 128

RESULT 15
ABX47313
ID ABX47313 standard; cDNA; 434 BP.
XX
AC ABX47313;

XX 21-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #12478.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-0960352.
XX
PR 12-JAN-1999; 99US-115707P.
PR 11-JAN-2000; 2000US-0480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
DR WPI; 2003-110599/10.
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and
PT analysis, cattle breeding, or for genetically improving cattle -
XX
PS Claim 2; SEQ ID No 12478; 245pp; English.
XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived
CC from cattle, and the LMFD nucleic acid can specifically hybridise to a
CC second nucleic acid molecule comprising any of 15112 nucleotide
CC sequences, appearing as ABX34836-ABX49947, or complements of them.
CC Also included are; (1) a transformed cell having a nucleic acid
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridisation between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMFD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome
CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMFD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139.
XX
SQ Sequence 434 BP; 135 A; 109 C; 86 G; 104 T; 0 other;

Query Match 92.1%; Score 74.6; DB 25; Length 434;
Best Local Similarity 95.1%; Pred. No. 2.4e-16;
Matches 77; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAAGCCCAACAATGTCCCTGCCACCCAGTGGCATGGCGTCT 60
|||||
Db 36 GCCACATCAGACTGGCTCAATGCCAACAATGTACCTGCTACCCAGTGGCATGGCCATCT 95

CC (CPSI) gene. The method is used to detect subjects at risk of
CC hepatitis, sclerosis, pulmonary hypertension and bone marrow
CC transplant toxicity. These conditions can be treated or prevented
CC by administration of a nitric oxide precursor or by gene
CC therapy (administration of sequences that encode CPSI).
XX
SQ Sequence 5761 BP; 1625 A; 1209 C; 1309 G; 1618 T; 0 other;

Query Match 98.0%; Score 79.4; DB 22; Length 5761;
Best Local Similarity 98.8%; Pred. No. 8.9e-18;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCAGTGGCGCTCT 60
Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCAGTGGCGCTCT 4359
|||||
QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 4360 CAAGAAGGACAGAATCCCAGC 4380
|||||

RESULT 12
AAC89487
ID AAC89487 standard; DNA; 5761 BP.
XX
AC AAC89487;
XX
DT 13-MAR-2001 (first entry)
XX Human carbamyl phosphate synthase DNA.
DE
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200073322-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15079.
XX
PR 01-JUN-1999; 99US-0323472.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Summar ML, Christman BW;
XX
DR WPI; 2001-049926/06.
XX
PT Detecting susceptibility to suboptimal urea cycle function, e.g. bone
PT marrow transplant toxicity, comprises identifying a polymorphism in the
PT gene for carbamylphosphate synthase -
XX
PS Claim 65; Page 146; 171pp; English.
XX
CC The present invention relates to screening for susceptibility to
CC suboptimal urea cycle function or to bone marrow transplant toxicity
CC by detecting a polymorphism in the carbamyl phosphate synthase I
CC (CPSI) gene. The method is used to detect subjects at risk of
CC hepatitis, sclerosis, pulmonary hypertension and bone marrow
CC transplant toxicity. These conditions can be treated or prevented
CC by administration of a nitric oxide precursor or by gene
CC therapy (administration of sequences that encode CPSI).
XX
SQ Sequence 5761 BP; 1625 A; 1209 C; 1309 G; 1618 T; 0 other;

Query Match 98.0%; Score 79.4; DB 22; Length 5761;
Best Local Similarity 98.8%; Pred. No. 8.9e-18;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCAGTGGCGCTCT 60
Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCAGTGGCGCTCT 4359
|||||
QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 4360 CAAGAAGGACAGAATCCCAGC 4380
|||||

RESULT 13
AAC89488
ID AAC89488 standard; DNA; 5761 BP.
XX
AC AAC89488;
XX
DT 13-MAR-2001 (first entry)
XX Human carbamyl phosphate synthase DNA.
DE
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200073322-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15079.
XX
PR 01-JUN-1999; 99US-0323472.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Summar ML, Christman BW;
XX
DR WPI; 2001-049926/06.
XX
PT Detecting susceptibility to suboptimal urea cycle function, e.g. bone
PT marrow transplant toxicity, comprises identifying a polymorphism in the
PT gene for carbamylphosphate synthase -
XX
PS Claim 65; Page 156-161; 171pp; English.
XX
CC The present invention relates to screening for susceptibility to
CC suboptimal urea cycle function or to bone marrow transplant toxicity
CC by detecting a polymorphism in the carbamyl phosphate synthase I
CC (CPSI) gene. The method is used to detect subjects at risk of
CC hepatitis, sclerosis, pulmonary hypertension and bone marrow
CC transplant toxicity. These conditions can be treated or prevented
CC by administration of a nitric oxide precursor or by gene
CC therapy (administration of sequences that encode CPSI).
XX
SQ Sequence 5761 BP; 1626 A; 1209 C; 1308 G; 1618 T; 0 other;

Query Match 98.0%; Score 79.4; DB 22; Length 5761;
Best Local Similarity 98.8%; Pred. No. 8.9e-18;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCAGTGGCGCTCT 60
Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCAGTGGCGCTCT 4359
|||||
QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 4360 CAAGAAGGACAGAATCCCAGC 4380
|||||

RESULT 14
ABZ84741
ID ABZ84741 standard; cDNA; 715 BP.
XX
AC ABZ84741;
XX
DT 14-MAY-2003 (first entry)

Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCAGTGGCGCTCT 4359
QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 4360 CAAGAAGGACAGAATCCCAGC 4380
|||||

RESULT 13
AAC89488
ID AAC89488 standard; DNA; 5761 BP.
XX
AC AAC89488;
XX
DT 13-MAR-2001 (first entry)
XX Human carbamyl phosphate synthase DNA.
DE
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200073322-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15079.
XX
PR 01-JUN-1999; 99US-0323472.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Summar ML, Christman BW;
XX
DR WPI; 2001-049926/06.
XX
PT Detecting susceptibility to suboptimal urea cycle function, e.g. bone
PT marrow transplant toxicity, comprises identifying a polymorphism in the
PT gene for carbamylphosphate synthase -
XX
PS Claim 65; Page 156-161; 171pp; English.
XX
CC The present invention relates to screening for susceptibility to
CC suboptimal urea cycle function or to bone marrow transplant toxicity
CC by detecting a polymorphism in the carbamyl phosphate synthase I
CC (CPSI) gene. The method is used to detect subjects at risk of
CC hepatitis, sclerosis, pulmonary hypertension and bone marrow
CC transplant toxicity. These conditions can be treated or prevented
CC by administration of a nitric oxide precursor or by gene
CC therapy (administration of sequences that encode CPSI).
XX
SQ Sequence 5761 BP; 1626 A; 1209 C; 1308 G; 1618 T; 0 other;

Query Match 98.0%; Score 79.4; DB 22; Length 5761;
Best Local Similarity 98.8%; Pred. No. 8.9e-18;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCAGTGGCGCTCT 60
Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCAGTGGCGCTCT 4359
|||||
QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 4360 CAAGAAGGACAGAATCCCAGC 4380
|||||

RESULT 14
ABZ84741
ID ABZ84741 standard; cDNA; 715 BP.
XX
AC ABZ84741;
XX
DT 14-MAY-2003 (first entry)

PF 09-MAY-1991; 91JP-0135902.
XX
PR 09-MAY-1991; 91JP-0135902.
XX
PA (HARA/) HARAGUCHI Y.
PA (MATS/) MATSUDA I.
PA (MORI/) MORI M.
XX
DR WPI; 1993-006237/01.
DR P-PSDB; AAR30636.
XX
XX Carbamyl phosphate synthetase I gene - used to detect diseases
PT caused by carbamyl phosphatase synthetase by using overlapped
PT cDNA contg. total human cDNA sequence
XX
PS Claim 1; Page 6-8; 12pp; Japanese.
XX
CC This sequence is the human carbamyl phosphate synthetase I (CPSI)
CC gene. This sequence or fragments of it may be used as probes to
CC detect mutations in the CPSI gene. CPSI deficiency disease may be
CC detected by using overlapping cDNA representing the full length cDNA
CC sequence of human CPSI. This sequence was isolated using three
CC amplified fragments derived from protein coding regions of the rat
CC CPSI cDNA as probes.
XX
SQ Sequence 5215 BP; 1484 A; 1104 C; 1205 G; 1422 T; 0 other;

Query Match 98.0%; Score 79.4; DB 14; Length 5215;
Best Local Similarity 98.8%; Pred. No. 8.7e-18;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCCACTGGCATGGCGTCT 60
|||||
Db 4295 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCCACTGGCATGGCGTCT 4354

QY 61 CAAGAAGGACAGAAATCCCAGC 81
|||||
Db 4355 CAAGAAGGACAGAAATCCCAGC 4375

RESULT 10
ABN95271
ID ABN95271 standard; DNA; 5215 BP.
XX
AC ABN95271;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #1769 used to diagnose liver cancer.
DE
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a

PT liver tissue sample -
XX
PS Claim 1; SEQ ID NO 1769; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 5215 BP; 1484 A; 1102 C; 1207 G; 1422 T; 0 other;

Query Match 98.0%; Score 79.4; DB 24; Length 5215;
Best Local Similarity 98.8%; Pred. No. 8.7e-18;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCCACTGGCATGGCGTCT 60
|||||
Db 4295 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCCACTGGCATGGCGTCT 4354

QY 61 CAAGAAGGACAGAAATCCCAGC 81
|||||
Db 4355 CAAGAAGGACAGAAATCCCAGC 4375

RESULT 11
AAC89479
ID AAC89479 standard; DNA; 5761 BP.
XX
AC AAC89479;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human N1405 carbamyl phosphate synthase I DNA.
DE
XX
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200073322-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15079.
XX
PR 01-JUN-1999; 99US-0323472.
XX
PA (UIVA-) UNIV VANDERBILT.
XX
PI Summar ML, Christman BW;
XX
DR WPI; 2001-049926/06.
XX
PT Detecting susceptibility to suboptimal urea cycle function, e.g. bone
PT marrow transplant toxicity, comprises identifying a polymorphism in the
PT gene for carbamylphosphate synthase -
XX
PS Disclosure; Page 124-130; 171pp; English.
XX
CC The present invention relates to screening for susceptibility to
CC suboptimal urea cycle function or to bone marrow transplant toxicity
CC by detecting a polymorphism in the carbamyl phosphate synthase I

OS Homo sapiens.
XX
PN WO200073322-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15079.
XX
PR 01-JUN-1999; 99US-0323472.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Summar ML, Christman BW;
XX
XX WPI; 2001-049926/06.
XX
PT Detecting susceptibility to suboptimal urea cycle function, e.g. bone
PT marrow transplant toxicity, comprises identifying a polymorphism in the
PT gene for carbamylphosphate synthase .
XX
PS Disclosure; Page 134-140; 171pp; English.
XX
CC The present invention relates to screening for susceptibility to
CC suboptimal urea cycle function or to bone marrow transplant toxicity
CC by detecting a polymorphism in the carbamyl phosphate synthase I
CC (CPSI) gene. The method is used to detect subjects at risk of
CC hepatitis, sclerosis, pulmonary hypertension and bone marrow
CC transplant toxicity. These conditions can be treated or prevented
CC by administration of a nitric oxide precursor or by gene
CC therapy (administration of sequences that encode CPSI).
XX
SQ Sequence 5761 BP; 1624 A; 1210 C; 1309 G; 1618 T; 0 other;

Query Match 100.0%; Score 81; DB 22; Length 5761;
Best Local Similarity 100.0%; Pred. No. 2.5e-18;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAGCGCCCAACATGTCCTGCCACCCAGTGGCGGTCT 60
Db 4300 GCCACATCAGACTGGCTCAGCGCCCAACATGTCCTGCCACCCAGTGGCGGTCT 4359

QY 61 CAAGAAGGACAGAAATCCAGC 81
Db 4360 CAAGAAGGACAGAAATCCAGC 4380

RESULT 8
ID AB235442 standard; cDNA; 5772 BP.
XX
AC AB235442;
XX
DT 05-FEB-2003 (first entry)
XX
DE Human gene expression profile polynucleotide SEQ ID NO 553.
XX
KW Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
KW gene expression; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200274979-A2.
XX
PD 26-SEP-2002.
XX
PF 20-MAR-2002; 2002WO-US08456.
XX
PR 20-MAR-2001; 2001US-276947P.
XX
PA (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
XX

PI Wan J, Wang Y;
XX
DR WPI; 2002-740862/80.
XX
PT New gene expression profile generated from primary, endothelial,
PT epithelial, and muscle cell types, useful for identifying disease
PT pathologies involving alterations of gene expression, e.g. cancer
XX
PS Disclosure; Page 691-693; 850pp; English.
XX
CC The invention relates to a gene expression profile comprising one or more
CC genes (AB234889-AB235692) and generated from a cell type. The cell type
CC is a coronary artery endothelium, umbilical artery or vein endothelium,
CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
CC endothelium, myometrium microvascular endothelium, keratinocyte
CC epithelium, bronchial epithelium, mammary epithelium, prostate
CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
CC small airway epithelium, renal epithelium, umbilical artery smooth
CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
CC osteoblasts or prostate stromal cell. The gene expression profile is used
CC for determining the level of RNA expression for a sample, determining the
CC phenotype of a cell and distinguishing cell types. The gene or a protein
CC expression profile is useful in identifying disease pathologies
CC involving alterations of gene expression. The assessment of expression
CC profiles may provide meaningful information with respect to tumour type
CC and stage, treatment methods, and prognosis. The gene or protein
CC expression profile may also be used for creating microarrays. The
CC microarray is useful for genetic and physical mapping of genomes, DNA
CC sequencing, genetic or medical diagnosis, genotyping of organisms,
CC confirming cell or tissue identifications and in identifying promising
CC antibiotics, antiviral or antifungal agents.
XX
SQ Sequence 5772 BP; 1602 A; 1180 C; 1295 G; 1571 T; 124 other;

Query Match 100.0%; Score 81; DB 24; Length 5772;
Best Local Similarity 100.0%; Pred. No. 2.5e-18;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAAGCCCAACATGTCCTGCCACCCAGTGGCGGTCT 60
Db 4310 GCCACATCAGACTGGCTCAAGCCCAACATGTCCTGCCACCCAGTGGCGGTCT 4369

QY 61 CAAGAAGGACAGAAATCCAGC 81
Db 4370 CAAGAAGGACAGAAATCCAGC 4390

RESULT 9
AAQ34768
ID AAQ34768 standard; DNA; 5215 BP.
XX
AC AAQ34768;
XX
DT 06-MAY-1993 (first entry)
XX
DE hCPSI gene.
XX
KW Human; carbamyl phosphate synthetase I; CPSI; probe; mutation;
KW CPSI deficiency disease; rat; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 119..4621
FT /*tag= a
XX
PN JP04335889-A.
XX
PD 24-NOV-1992.
XX

XX Lung cancer; cytostatic; vaccine; gene therapy; cancer;
KW gene; ss.

XX Homo sapiens.

PN US2002172952-A1.

XX 21-NOV-2002.

XX 10-JUL-2001; 2001US-0902941.

XX 30-JUN-1999; 99US-0346492.

PR 15-OCT-1999; 99US-0419356.

PR 17-DEC-1999; 99US-0466867.

PR 30-DEC-1999; 99US-0476300.

PR 06-MAR-2000; 2000US-0519642.

PR 22-MAR-2000; 2000US-0533077.

PR 10-APR-2000; 2000US-0546259.

PR 27-APR-2000; 2000US-0560406.

PR 05-JUN-2000; 2000US-0589184.

PR 11-JUL-2000; 2000US-0614124.

PR 29-AUG-2000; 2000US-0651563.

PR 08-SEP-2000; 2000US-0658824.

PR 26-SEP-2000; 2000US-0671325.

PR 06-OCT-2000; 2000US-0677419.

PR 30-OCT-2000; 2000US-0702705.

PR 13-DEC-2000; 2000US-0736457.

PR 03-MAY-2001; 2001US-0849626.

XX (CORI-) CORIXA CORP.

PA Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW, Durham M;

PI Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;

XX WPI; 2003-328427/31.

XX New polynucleotide, useful for preparing a composition for treating or

PT inhibiting development of cancer, e.g. lung cancer -

XX Example 1; SEQ ID NO 93; 82pp; English.

XX The invention describes an isolated polynucleotide comprising one of 32

CC sequences, complement or degenerate variants of them. The polynucleotide

CC is useful for preparing a composition e.g. a vaccine or for gene therapy,

CC for treating or inhibiting development of cancer, e.g. lung cancer.

CC This sequence represents a polynucleotide associated with the

CC compositions and methods for the therapy and diagnosis of lung cancer.

XX Sequence 531 BP; 157 A; 139 C; 108 G; 127 T; 0 other;

QY Query Match 100.0%; Score 81; DB 25; Length 531;

Db Best Local Similarity 100.0%; Pred. No. 1.5e-18;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60

Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 216

QY 61 CAAGAAGGACAGATCCAGC 81

Db 217 CAAGAAGGACAGATCCAGC 237

RESULT 6

AAH57467

ID AAH57467 standard; cDNA; 5195 BP.

XX 10-SEP-2001 (first entry)

XX Human liver cell specific cDNA sequence SEQ ID NO:307.

XX Human; tissue specific; diagnosis; brain; heart; skeletal muscle;

KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;

KW metabolic disease; developmental disease; cytostatic; immunomodulatory;

KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.

XX Homo sapiens.

OS WO200132927-A2.

XX 10-MAY-2001.

XX 02-NOV-2000; 2000WO-US30396.

XX 04-NOV-1999; 99US-0163508.

XX (INCY-) INCYTE GENOMICS INC.

XX Sornasse T, Seilhamer JJ, Watson GA;

XX WPI; 2001-291057/30.

XX New cell and tissue specific polynucleotides useful for diagnosis,

XX prognosis or monitoring of treatments for disorders where the gene is

XX associated with a cancer, immunopathology or neuropathology -

XX Claim 1; Page 230-231; 327pp; English.

XX AAH57161 to AAH57576 represent cell and tissue specific polynucleotide

CC sequences (I). (I) can have cytostatic, immunomodulatory and

CC neuroprotective activities, and can be used in gene therapy. (I) and

CC proteins (II) encoded by then are used in high throughput screening

CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,

CC mimetics, peptides, proteins, agonists, antagonists, antibodies or

CC their fragments, immunoglobulins, inhibitors, drug compounds and

CC pharmaceutical agents. Expression of (I) in a sample indicates the

CC differentiation of embryonic stem cells into a tissue selected from

CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic

CC tissues. (I) and (II) are used to produce an expression profile that

CC defines a metabolic or developmental process, treatment, condition,

CC disease or disorder. The gene profile can be used for diagnosis,

CC prognosis or monitoring of treatments and for investigating a

CC predisposition to a disorder where the gene is associated with a

CC cancer, immunopathology or neuropathology.

XX Sequence 5195 BP; 1469 A; 1102 C; 1211 G; 1413 T; 0 other;

QY Query Match 100.0%; Score 81; DB 22; Length 5195;

Db Best Local Similarity 100.0%; Pred. No. 2.4e-18;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60

Db 4310 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 4369

QY 61 CAAGAAGGACAGATCCAGC 81

Db 4370 CAAGAAGGACAGATCCAGC 4390

RESULT 7

AAH57480

ID AAC89480 standard; DNA; 5761 BP.

XX AAC89480;

XX 13-MAR-2001 (first entry)

XX Human T1405 carbamyl phosphate synthase I DNA.

XX Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;

XX sclerosis; pulmonary hypertension; bone marrow; gene therapy; ds.

XX WPI; 2002-164634/21.
XX Novel polynucleotide encoding a lung tumour polypeptide useful for
PT stimulating and/or expanding T cells specific for a tumour protein
XX
PS Example 1; SEQ ID No 93; 223pp; English.
XX
CC The invention describes an isolated polynucleotide and polypeptide
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein for determining the presence of a cancer in a patient. A
CC composition containing the polynucleotide and/or polypeptide is useful
CC for treating a lung cancer in a patient. The polypeptide is useful for
CC removing tumour cells from a biological sample. The polynucleotide is
CC also useful as probe or primer to detect the level of mRNA encoding a
CC tumour protein. This sequence encodes a lung tumour associated protein
CC or protein fragment, described in the method of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 531 BP; 157 A; 139 C; 108 G; 127 T; 0 other;

Query Match 100.0%; Score 81; DB 24; Length 531;
Best Local Similarity 100.0%; Pred. No. 1.5e-18;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCGTCT 60
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCGTCT 216

QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 217 CAAGAAGGACAGAATCCCAGC 237

RESULT 4
ACA10415
ID ACA10415 standard; cdna; 531 BP.
XX ACA10415;
AC
DT 05-JUN-2003 (first entry)
XX
DE Human lung cancer-associated cdna, SEQ ID 93.
XX
KW Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW T cell expansion; CD4; CD8.
XX
OS Homo sapiens.
XX
PN US2002197669-A1.
XX
PD 26-DEC-2002.
XX
PF 03-MAY-2001; 2001US-0849626.
XX
PR 13-DEC-2000; 2000US-0736457.
XX
PA (BANG/) BANGUR C S.
PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.
XX
PI Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;
PI Clapper JD;
XX
DR WPI; 2003-352750/33.
XX

PT Novel lung cancer polynucleotide encoding lung cancer protein, useful
PT for detecting the presence of lung cancer in a patient, and in
PT pharmaceutical compositions, e.g. vaccines, for treating lung cancer
XX
PS Example 1; Page -: 72pp; English.
XX
CC The invention relates to a polynucleotide encoding a lung tumour protein,
CC comprising a sequence selected from any of the 14 sequences
CC mentioned in the specification, or a sequence (S2) mentioned in
CC specification, complement of S1, sequences consisting of at least 20
CC contiguous residues of S1, sequences that hybridise to S1, sequences
CC having 75%, preferably 90%, identity to S1, or degenerate variants of
CC S1. Also included are an isolated polypeptide (comprising a sequence (S3)
CC selected from any one of the 4 amino acid sequences mentioned in the
CC specification, a sequence encoded by the polynucleotide, or sequences
CC having at least 70%, preferably 90%, identity to a sequence encoded by
CC the polynucleotide), an expression vector comprising the polynucleotide
CC operably linked to an expression control sequence, a host cell
CC transformed or transfected with the vector, an isolated antibody (or its
CC antigen-binding fragment) that specifically binds to the polypeptide,
CC detecting the presence of a cancer in a patient, a fusion protein
CC comprising the polypeptide, an oligonucleotide that hybridises to
CC S1 under moderately stringent conditions, stimulating and/or expanding T
CC cells specific for a tumour protein (comprising contacting T cells with
CC the polynucleotide, protein or antigen-presenting cells, under conditions
CC and for a time sufficient to permit the stimulation and/or expansion of T
CC cells) and inhibiting the development of a cancer in a patient (by
CC incubating CD4+ and/or CD8+ T cells isolated from a patient with the
CC polynucleotide, protein or antigen presenting cells that express the
CC polynucleotide, such that T cells proliferate, administering to the
CC patient an effective amount of the proliferated T cells, and thus
CC inhibiting the development of a cancer in the patient. The
CC polynucleotide, protein ands cells are useful in a composition for
CC stimulating an immune response in a patient, and for treating a cancer in
CC a patient (particularly lung cancer). The oligonucleotide is useful for
CC determining the presence of a cancer in a patient. The protein and
CC oligonucleotides are useful in pharmaceutical compositions, e.g.
CC vaccines. The polynucleotide is also useful as a probe or primer for
CC nucleic acid hybridisation, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. An amplified portion of the polynucleotide is
CC useful for isolating a full-length gene from a suitable library.
CC The present sequence is a cdna (full length, extended or partial)
CC isolated from a library derived from lung tumour/cancer cells.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from the USPTO
CC at seqdata.uspto.gov/sequence.html?DocId=20020197669.
XX
SQ Sequence 531 BP; 157 A; 139 C; 108 G; 127 T; 0 other;

Query Match 100.0%; Score 81; DB 25; Length 531;
Best Local Similarity 100.0%; Pred. No. 1.5e-18;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCGTCT 60
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCGTCT 216

QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 217 CAAGAAGGACAGAATCCCAGC 237

RESULT 5
ABX99366
ID ABX99366 standard; cdna; 531 BP.
XX
AC ABX99366;
XX
DT 22-MAY-2003 (first entry)
XX
DE Lung cancer therapyand diagnosis associated cdna #93.

PT gene for carbamylphosphate synthase -
XX Disclosure; Fig 10; 171pp; English.
PS
XX
CC The present invention relates to screening for susceptibility to
CC suboptimal urea cycle function or to bone marrow transplant toxicity
CC by detecting a polymorphism in the carbamyl phosphate synthase I
CC (CPSI) gene. The method is used to detect subjects at risk of
CC hepatitis, sclerosis, pulmonary hypertension and bone marrow
CC transplant toxicity. These conditions can be treated or prevented
CC by administration of a nitric oxide precursor or by gene
CC therapy (administration of sequences that encode CPSI).
XX
SQ Sequence 495 BP; 158 A; 96 C; 85 G; 155 T; 1 other;

Query Match 100.0%; Score 81; DB 22; Length 495;
Best Local Similarity 100.0%; Pred. No. 1.4e-18;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAGCGCCCAACAATGTCCCTGCCACCCAGTGGCATGGCGGTCT 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
126 GCCACATCAGACTGGCTCAGCGCCCAACAATGTCCCTGCCACCCAGTGGCATGGCGGTCT 185

QY 61 CAGCAAGGACAGAAATCCCAGC 81
Db ||||||||||||||||||||||||||||
186 CAGCAAGGACAGAAATCCCAGC 206

RESULT 2
AAF68175
ID AAF68175 standard; cDNA; 531 BP.
XX
AC AAF68175;
XX
DT 12-APR-2001 (first entry)
XX
DE Human lung tumour protein related nucleotide sequence SEQ ID NO:93.
XX
KW Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
KW cytostatic; antisense inhibition; ss.
XX
OS Homo sapiens.
XX
PN WO200100828-A2.
XX
PD 04-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18061.
XX
PR 30-JUN-1999; 99US-0346492.
PR 15-OCT-1999; 99US-0419356.
PR 17-DEC-1999; 99US-0466867.
PR 30-DEC-1999; 99US-0476300.
PR 06-MAR-2000; 2000US-0519642.
PR 22-MAR-2000; 2000US-0533077.
PR 10-APR-2000; 2000US-0546259.
PR 27-APR-2000; 2000US-0560406.
PR 05-JUN-2000; 2000US-0589184.
XX
PA (CORI-) CORIXA CORP.
XX
XX Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
PI Retter MW, Mannion J;
PI
XX WPI; 2001-071488/08.
XX
PT Lung tumor-associated proteins and the nucleic acids that encode them,
PT useful for preventing, diagnosing and treating lung cancer -
XX
PS Example 1; Page 181; 436pp; English.
XX
CC The present invention describes immunogenic portions of lung tumour-

CC associated proteins (I) and the nucleic acids (NAs) that encode them.
CC (I) have cytostatic activity and can be used in gene therapy, antisense
CC inhibition and in vaccines. The NAs and the lung tumour-associated
CC proteins they encode may be used in the prevention, treatment and
CC diagnosis of diseases associated with their inappropriate expression,
CC especially lung cancers. For example, the NAs may be administered to
CC treat diseases by rectifying mutations or deletions in a patient's genome
CC that affect the activity of the protein by expressing inactive proteins
CC or to supplement the patient's own production of (I). Additionally, the
CC NAs may be used to produce the lung-tumour associated protein, according
CC to standard recombinant DNA methodology. Conversely, antisense NA
CC molecules may be administered to down regulate protein expression by
CC binding with the cells own genes and preventing their expression. The NA
CC and complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar NA sequences in
CC samples, and hence which patients may be in need of treatment for lung
CC cancer. The (I) may be used as antigens in the production of antibodies
CC and in assays to identify modulators (agonists and antagonists) of the
CC expression and activity of the protein. AAF68083 to AAF68878 and
CC AAB76848 to AAB76878 represent human lung tumour protein related
CC nucleotide and protein sequences which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 531 BP; 157 A; 139 C; 108 G; 127 T; 0 other;

Query Match 100.0%; Score 81; DB 22; Length 531;
Best Local Similarity 100.0%; Pred. No. 1.5e-18;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGTCAACGCCAACAAATGTCCCTGCCACCCAGTGGCATGGCGGTCT 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
157 GCCACATCAGACTGGTCAACGCCAACAAATGTCCCTGCCACCCAGTGGCATGGCGGTCT 216

QY 61 CAGCAAGGACAGAAATCCCAGC 81
Db ||||||||||||||||||||||||
217 CAGCAAGGACAGAAATCCCAGC 237

RESULT 3
ABK38086
ID ABK38086 standard; cDNA; 531 BP.
XX
AC ABK38086;
XX
DT 21-MAY-2002 (first entry)
XX
DE cDNA encoding clone #19065 of lung tumour protein.
XX
KW Lung tumour; cancer; T cell; immune response stimulator;
KW cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200204514-A2.
XX
PD 17-JAN-2002.
XX
PF 10-JUL-2001; 2001WO-US22058.
XX
PR 11-JUL-2000; 2000US-0614124.
PR 29-AUG-2000; 2000US-0651563.
PR 08-SEP-2000; 2000US-0658824.
PR 26-SEP-2000; 2000US-0671325.
PR 06-OCT-2000; 2000US-0677419.
PR 30-OCT-2000; 2000US-0702705.
PR 13-DEC-2000; 2000US-0736457.
PR 03-MAY-2001; 2001US-0849626.
XX
PA (CORI-) CORIXA CORP.
XX
XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;
PI McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 22:19:36 ; Search time 201.5 seconds
(without alignments)
1085.134 Million cell updates/sec

Title: US-09-585-077C-3_COPY_4300_4380
Perfect score: 81
Sequence: 1 gccacatcagactggctcaa.....aagaaggacagaatcccagc 81

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	495	22 AAC89481	Target 5' and 3' s
2	81	100.0	531	22 AAF68175	Human lung tumour
3	81	100.0	531	24 ABK38086	cDNA encoding clon
4	81	100.0	531	25 ACA10415	Human lung cancer-
5	81	100.0	531	25 ABX93366	Lung cancer therap
6	81	100.0	5195	22 AAH57467	Human liver cell s
7	81	100.0	5761	22 AAC89480	Human T1405 carbam
8	81	100.0	5772	24 ABZ35442	Human gene express

9	79.4	98.0	5215	14	AAQ34768	hCPSI gene. Homo
10	79.4	98.0	5215	24	ABN95271	Gene #1769 used to
11	79.4	98.0	5761	22	AAC89479	Human N1405 carbam
12	79.4	98.0	5761	22	AAC89487	Human carbamyl pho
13	79.4	98.0	5761	22	AAC89488	Human carbamyl pho
14	76.8	94.8	715	25	ABZ84741	Toxicologically re
15	74.6	92.1	434	25	ABX47313	Bovine EST associa
16	67.4	83.2	816	24	ABA93013	Human cDNA clone B
17	64.4	79.5	459	22	AAH57259	Human liver specif
18	47.6	58.8	580	24	ABT08945	Phase-1 Rat CT gen
19	41.8	51.6	242	25	ABX36653	Bovine EST associa
20	29.4	36.3	31	22	AAI31280	Human single nucle
21	29.4	36.3	3591	22	AAI02819	Human reproductive
22	28	34.6	2153	24	ABI99890	Mouse ischaemic co
23	28	34.6	2168	18	AAV00499	Murine alpha-(2) s
24	27.4	33.8	646	24	ABT09297	Phase-1 Rat CT gen
25	27	33.3	288	25	ABX54929	Bovine EST associa
26	26.4	32.6	969	16	AAQ87619	Morchella sp. DNA
27	26.4	32.6	3276	16	AAQ87604	Fungus-infected G.
28	26.4	32.6	3276	16	AAQ88050	Glucan Lyase 1 gen
29	26.4	32.6	3276	20	AAV84193	Gracilariaopsis lem
30	26	32.1	1392	21	AAZ53795	Neisseria meningit
31	26	32.1	15896	21	AAH81517	N. meningitidis pa
32	26	32.1	34980	21	AAF21608	Neisseria meningit
33	26	32.1	34980	21	AAF21609	Neisseria meningit
34	26	32.1	1437668	21	AAH81490	N. meningitidis B
35	25.6	31.6	1202	19	AAV71981	A nucleotide seque
36	25.6	31.6	3345	24	ABK84183	Human cDNA differe
37	25.4	31.4	1703	21	AAZ90670	Human KCNE1 protei
38	25.4	31.4	1703	21	AAZ98902	Human long QT synd
39	25.4	31.4	1703	22	AAC89983	Human KCNE1 coding
40	25.2	31.1	284	21	AAC05908	Human secreted pro
41	25.2	31.1	8098	22	AAK90713	Human digestive sy
42	25.2	31.1	9556	22	AAK90714	Human digestive sy
43	25	30.9	25	25	ABZ84401	Toxicologically re
44	25	30.9	450	24	ABK78235	Bacillus clausii g
45	25	30.9	2372	8	AAH70814	Aspergillus niger

ALIGNMENTS

RESULT 1	
AAC89481	
ID AAC89481 standard; DNA; 495 BP.	
XX	
AC AAC89481;	
XX	
DT 13-MAR-2001 (first entry)	
DE Target 5' and 3' sequence.	
XX	
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;	
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy; ds.	
XX	
OS Homo sapiens.	
XX	
PN WO200073322-A1.	
XX	
PD 07-DEC-2000.	
XX	
PF 01-JUN-2000; 2000WO-US15079.	
XX	
PR 01-JUN-1999; 99US-0323472.	
XX	
PA (UYVA-) UNIV VANDERBILT.	
XX	
PI Summar ML, Christman BW;	
XX	
DR WPI; 2001-049926/06.	
XX	
PT Detecting susceptibility to suboptimal urea cycle function, e.g. bone marrow transplant toxicity, comprises identifying a polymorphism in the	

TAHIVLEDTKMGYSFGHPSSVAGEVVENTGLGGYPEAITDPAYKGQILTMANPIIG
NGGAPDTSLSDELGLSKYLENSNGIKVSGLLVLDDYSKPDYNNHLATKSLGQWLQBEKVPA
IYGVDTRMLTKIIRDKGTMLGKIEFEGQPVDFVDPNKNLIAEYSTKDVKVYKGNPT
KVAVDCG1KNNVIRLLYKRGAEVHLVPWNHDFTRMEYDGLIAGGPGNPALEPLIQ
NVQKILESDEKPELFGISTGNLITGLAAGAKTYKMSMANRGQNPVLNITNKQAFITA
QNHCYALDNTPLPAGWKPLFVNVDQTNNEGIMHESKPFPAVQFHPEVTPGPIDEYLFED
SFFSLIKKGRATTITSVLPKPALVASRVEVSKVLILSGGSLIGQAGEFDYSGQAVK
AMKEENVKTVLMPNPIASQTVNEVGLQADTVYFLPITPQFVTEVIKAEQPPDGLILGM
GGOTALNCGVLEFKRGVLKEYGVKVLGTSVESIMATEDROLFSDKLINEINEKIAPSFA
VESIEDALKAADTIGYPVMIRSAYALGGGSGICPNRETLMDSLTKAFAMTNOILVEK
SVTGWKEIEYEVVRDADNCVTVCMENVDAMGVHTGDSVVVAPAQTLISNAEQMLRR
TSINVVHRHGIIVGECNIOFALHPTSMETCIIIEVNARSPNSALASKTTGYPLAFIAAK
IALGIPLPKIKNVVSGKTSACFEPSLDYMTVKIPRWDLDRFHTSSRIGSSMKSVEV
MAIGRTFEESFQKALRMCHPSIEGTPRLPMNKEWPSNLDLRKELSEPSSTRITAIK
AIDDNMSLEIEKLTYYIDKWFLYKMRDILNMEKTLKGLNSESMTETLKRAKEIGFSD
KOISKCLGTFEATRELRLKKNHHPWKQIDTTLAAYPSVTNLYVTYNGQEHVDNFD
DEGMVVLGCGYHIGSSVEFDWCAVSSIRTLRQLGKKTVVVNCNPEVTSTDFDECDKL
YFEELSLEIRLIDYHOEACGCIISVGGQIPNNLAVPLYKNGVKIMGTSPLIQIDRAED
RSIFSALVDELKVAQAPWKAVNTLNEALEFAKSDYDPCLLRPSYVLSGSAWNVFSED
EMKKFLEEATRVSOATPVVLTKFVEGAREVEMDAVGKDRVISHAISEHVEDAGVHSE
NATLMLPTQITISQGAIEKVKDATRKAFAISGPFNVQFLVKGNDVLVNECNLRASR
SFFSVSKTLGVDFIDVATKVLIGENVDEKHLPTLDHPIDYVVAIKAPMFSWPERLRD
ADPILRCEMASTGEVACFGEIHHTAFLKAMLSFGKIPQKILIGIQQSFRPRFLGVA
EQLHNEGFKLEATEATSDWLNANNVPANVPWPSQEGQNPSSLIRKLIRDGSDILVI
NLPNNNTKFFVDNVVIRRTAVDSGIPLLTNEQVTKLFAEAVQKSRKRVDSKSLFHYROY
SAGKAA"

BASE COUNT 1484 a 1102 c 1207 g 1422 t
ORIGIN

Query Match 98.0%; Score 79.4; DB 9; Length 5215;
Best Local Similarity 98.8%; Pred. No. 7.6e-16;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
|||||
Db 4295 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 4354
|||||

QY 61 CAAGAAGGACAGAATCCCAGC 81
|||||

Db 4355 CAAGAAGGACAGAATCCCAGC 4375
|||||

RESULT 15
AR184416
LOCUS AR184416 5761 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 1 from patent US 6346382.
ACCESSION AR184416
VERSION AR184416.1 GI:20230381
KEYWORDS
SOURCE . Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5761)
AUTHORS Summar,M.L. and Christman,B.W.
TITLE Human carbamyl phosphate synthetase I polymorphism and diagnostic methods related thereto
JOURNAL Patent: US 6346382-A 1 12-FEB-2002;
FEATURES Location/Qualifiers
source
1..5761
/organism="unknown"
BASE COUNT 1625 a 1209 c 1309 g 1618 t
ORIGIN

Query Match 98.0%; Score 79.4; DB 6; Length 5761;
Best Local Similarity 98.8%; Pred. No. 7.4e-16;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
|||||

Db 4300 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 4359
|||||

QY 61 CAAGAAGGACAGAATCCCAGC 81
|||||

Db 4360 CAAGAAGGACAGAATCCCAGC 4380

Search completed: September 12, 2003, 23:21:52
Job time : 1418.5 secs

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGCTCT 60
|||||
Db 30750 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGCTCT 30691

QY 61 CAAGAAGGACAGATCCAGC 81
|||||
Db 30690 CAAGAAGGACAGATCCAGC 30670

RESULT 12
AX409122
LOCUS AX409122 5215 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 1769 from Patent WO0229103.
ACCESSION AX409122
VERSION AX409122.1 GI:21441827
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 1769 11-APR-2002;
GENE LOGIC INC (US)

FEATURES
source Location/Qualifiers
1..5215
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. D90282"

BASE COUNT 1484 a 1102 c 1207 g 1422 t
ORIGIN

Query Match 98.0%; Score 79.4; DB 6; Length 5215;
Best Local Similarity 98.8%; Pred. No. 7.6e-16;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGCTCT 60
|||||
Db 4295 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGCTCT 4354

QY 61 CAAGAAGGACAGATCCAGC 81
|||||
Db 4355 CAAGAAGGACAGATCCAGC 4375

RESULT 13
E04065
LOCUS E04065 5215 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding human carbamyl phosphate synthetase 1.
ACCESSION E04065
VERSION E04065.1 GI:2172275
KEYWORDS JP 1992335889-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 5215)
AUTHORS Haraguchi,Y., Mori,M. and Matsuda,I.
TITLE CARBAMYLPHOSPHATE SYNTHETASE I GENE, DETECTION OF MUTATION OF THE
GENE AND DNA PROVE TO USED THEREFOR
JOURNAL Patent: JP 1992335889-A 1 24-NOV-1992;
HARAGUCHI YOGO, MORI MASATAKA, MATSUDA ICHIRO
COMMENT OS Homo sapiens (human)
PN JP 1992335889-A/1
PD 24-NOV-1992
PF 09-MAY-1991 JP 1991135902
PI HARAGUCHI YOGO, MORI MASATAKA, MATSUDA ICHIRO PC
C12N15/52,C07H21/04,C12N15/11,C12Q1/68//A61B10/00; CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;

CC anti-sense: No;
CC *source: tissue=Liver;
FH Key Location/Qualifiers
FH 5'UTR 1..118
FT CDS 119..4621
FT /product='carbamyl phosphate synthetase 1' FT
3'UTR 4622..5215.

FEATURES
source Location/Qualifiers
1..5215
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"
BASE COUNT 1484 a 1102 c 1207 g 1422 t
ORIGIN

Query Match 98.0%; Score 79.4; DB 6; Length 5215;
Best Local Similarity 98.8%; Pred. No. 7.6e-16;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGCTCT 60
|||||
Db 4295 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGCTCT 4354

QY 61 CAAGAAGGACAGATCCAGC 81
|||||
Db 4355 CAAGAAGGACAGATCCAGC 4375

RESULT 14
HUMCP51
LOCUS HUMCP51 5215 bp mRNA linear PRI 29-MAY-2002
DEFINITION Homo sapiens mRNA for carbamyl phosphate synthetase I, complete
cds.
ACCESSION D90282
VERSION D90282.1 GI:219552
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 5215)
AUTHORS Haraguchi,Y., Uchino,T., Takiguchi,M., Endo,F., Mori,M. and
Matsuda,I.
TITLE Cloning and sequence of a cDNA encoding human carbamyl phosphate
synthetase I: molecular analysis of hyperammonemia
JOURNAL Gene 107 (2), 335-340 (1991)
MEDLINE 92084128
PUBMED 1840546
COMMENT Submitted (18-JAN-1991) to DDBJ by:
Yougo Haraguchi
Department of Pediatrics
Kumamoto University Medical School
1-1-1 Honjo
Kumamoto 860
Japan
Phone: 096-344-2111 x5654
Fax: 096-366-3471.

FEATURES
source Location/Qualifiers
1..5215
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/tissue_type="liver"
/note="short arm"
119..4621
/EC_number="6.3.4.16"
/codon_start=1
/product="carbamyl phosphate synthetase I"
/protein_id="BAA14328.1"
/db_xref="GI:219553"
/translation="MTRLTAPKVVVRLTKTGFGFTNVTAHQKWKFSRPGIRLLSVKQ"

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Rateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-430P16, 200 base pair overlap; the clone sequenced to the right is RP11-485G2, 200 base pair overlap. Actual start of this clone is at base position 195 of RP11-349G4; actual end is at base position 183502 of RP11-349G4.

RP11-349G4 contains a transposon which has been omitted from the submitted sequence. The transposon would insert after base position 170720 of this sequence.

FEATURES

source

1. .183696
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-349G4"
/clone_lib="RPCI-11"

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region
13293. .13847
/rpt_family="L1"
repeat_region
15165. .15620
/rpt_family="L2"
repeat_region
16293. .16360
/rpt_family="L2"
repeat_region
17745. .17990
/rpt_family="MIR"
repeat_region
18558. .19067
/rpt_family="MaLR"
repeat_region
19260. .19646
/rpt_family="MaLR"
repeat_region
19955. .20104
/rpt_family="MIR"
repeat_region
20688. .20730
/rpt_family="L2"
repeat_region
20744. .20983
/rpt_family="MIR"
repeat_region
21292. .21726
/rpt_family="L2"
repeat_region
23120. .23421
/rpt_family="Alu"
repeat_region
26196. .26367
/rpt_family="MIR"
repeat_region
26530. .26758
/rpt_family="Alu"
misc_feature
27385. .27943
/note="similar to EST N58033 (NID:g1201923) yv63d03.s1"
misc_feature
27386. .28085
/note="similar to EST AI936814 (NID:g5675684) wp69g09.x1"
misc_feature
27386. .28068
/note="similar to EST AA676489 (NID:g2657011) ad39d05.s1"
misc_feature
27386. .27880
/note="similar to EST AAL79297 (NID:gl760649) zp45b07.s1"
misc_feature
27386. .27823
/note="similar to EST AV646249 (NID:g9867263) "
misc_feature
27386. .27819
/note="similar to EST AA113231 (NID:gl664935) zn66c01.s1"
misc_feature
27386. .27743
/note="similar to EST AV645723 (NID:g9866737) "
misc_feature
27386. .27587
/note="similar to EST AW235498 (NID:g6567887) xn13f11.x1"
misc_feature
27387. .27904
/note="similar to EST AA777572 (NID:g2837051) zi98b01.s1"
misc_feature
27387. .27814
/note="similar to EST AA191429 (NID:gl780108) zp83g12.s1"
misc_feature
27388. .27612
/note="similar to EST N54164 (NID:g1195330) yv64b09.s1"
misc_feature
27389. .27736
/note="similar to EST T52465 (NID:g654325) yb32b06.s1"
misc_feature
27391. .28148
/note="similar to EST AI132943 (NID:g6360259) "
misc_feature
27392. .27724
/note="similar to EST AI196432 (NID:g3749038) ui69f12.x1"
misc_feature
27407. .27724
/note="similar to EST T69421 (NID:g680569) yc38d06.s1"
misc_feature
27486. .28021
/note="similar to EST AV650985 (NID:g9871999) "
misc_feature
27487. .28199
/note="similar to EST AV650994 (NID:g9872008) "
misc_feature
27488. .28121
/note="similar to EST AV661872 (NID:g9882886) "
misc_feature
27513. .28021
/note="similar to EST AV651545 (NID:g9872559) "
misc_feature
27519. .28148
/note="similar to EST AF074674 (NID:g6470492) "
misc_feature
27582. .28199

Query Match

Best Local Similarity 100.0%; Score 81; DB 9; Length 183696;

Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

REFERENCE 1 (bases 1 to 121959)
AUTHORS Haberle,J. and Koch,H.G.
TITLE Human carbamylphosphate synthetase 1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 121959)
AUTHORS Haberle,J. and Koch,H.G.
TITLE Direct Submission
JOURNAL Submitted (08-AUG-2002) Pediatrics, University of Munster,
Albert-Schweitzer-Str. 33, Munster 48129, Germany
FEATURES
source
1. 121959
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q35"
join(<116..241,16680..16789,19728..19872,20803..20892,
23096..23152,25999..26091,31440..31529,33488..33616,
34182..34288,35213..35351,36261..36338,37890..37988,
38869..38964,42754..42943,43937..44094,45584..45712,
48484..48628,50113..50323,51743..51941,55499..55675,
59805..59923,81084..81225,82532..82597,83378..83441,
85866..86047,91245..91439,91855..91922,93745..93820,
97407..97484,99907..100014,101981..102070,103867..104037,
106505..106579,111568..111666,118284..118343,
119110..119222,120389..120518,121269..>121367)
/product="carbamylphosphate synthetase 1"
join(116..241,16680..16789,19728..19872,20803..20892,
23096..23152,25999..26091,31440..31529,33488..33616,
34182..34288,35213..35351,36261..36338,37890..37988,
38869..38964,42754..42943,43937..44094,45584..45712,
48484..48628,50113..50323,51743..51941,55499..55675,
59805..59923,81084..81225,82532..82597,83378..83441,
85866..86047,91245..91439,91855..91922,93745..93820,
97407..97484,99907..100014,101981..102070,103867..104037,
106505..106579,111568..111666,118284..118343,
119110..119222,120389..120518,121269..121367)
/EC_number="6.3.4.16"
/codon_start=1
/product="carbamylphosphate synthetase 1"
/protein_id="AAN7181.1"
/db_xref="GI:25992606"
/translation="MTRLIAFKVVRTLKTGFGFTNTVTAHQKWKFSRPGIRLLSVKAO
TAHIVLEDGTRKMGYSFGHPSSVAGEVVENTGLGGYPEAITDPAYKGQILTMANPIIG
NGAPDTTALDELGSKYLENGIKVSGLLVLDYDKDYNHNLATKSLGQWLQEEKVP
IYGVDTMLTKLIRDKGTMLGKIEFGQPVDFVDPNKONLIAEVSTKDKVYKGGNPT
KVVAVDCGIKNNVIRLLVVKRAEVHLYPNWHDFTKMEYDGLIACGCPNPAELLIQ
NVKILESDRKEPLFGISTGNLITGLAAGAKTYKMSMANRGQNPVLNITNKQAFITA
QNHGYALDNLTPAGWKPLFVNVDQTNEGIMHESKPFPAVFQHPFVTPGPDITVELFD
SFTSLIKKGKATITSVLPKPALVASRVEVSKVLILGSGLSIGQAGEFDYSGSOAVK
AMKEENVKTVLMNPNIASVQNEVGLKQADTVYFLPIPPQFVTEVIKAEQPDGLIGM
GGOTALNCVELFKRGVLKEYGVKVLGTSVESIMATEDRQLFSDKLINEIKIAPSPA
VESTEDALKAADTIGPYMIRSAYALGGLSGICPNRETLMDLSTKAFAMTNQILVEK
SVTGWKEIEYEVVRDADONCVTVYCNMENVDMGVHTGDSVVVAPAOQLSNAEFQMLRR
TSINVRHLGIVGECNIQFALHPTSMEYCIIEVNARLSRSALASAKATGYPLAFTAAK
IALGIPLPEIKNVVSGKTSACFEPSLDYMTKIPRWDLDRFHGTSSRIGSSMKSVGEV
MAIGRTFEESQKALRMCHPSIEGFTPRLPNNKEWPSNLDLRKELSEPSSTRIYAIK
AIDNMSLDEIEKITYIDKWELYKMRDILNMEKTLKGLNSESMTTEETLKRAKEIGFSD
KQISKCLGLETAQTRELRLKKNHPWKQIDTLAAEYPSVTNLYLYPYNGQEHVDNFD
DHGMVILGCGPHIGSSVEFDCAVSSIRTLRQLGKKTVVNCNPEVSTDFDECCKL
YFELSRLERILDIYHQEACGGCIISVGQIPNNLAVPLYKNGVKIMGTSPLOIDRAED
RSIFSASVDELKVAQAPWKAVNTLNEALEAKSDVDPCLLRPSYVLSSGSMNVVFSGD
EMKFLLEEATRVQEPHVLKFFVEGAREVEMDVGKGRVISHAISEHVEDVGHSG
DATMLPTQTISOGAIEKVKDATRKIAKAFASISGPFNVQFLVKGNDVILVIECNLRASR
SEPFVSKTLGVDFIDVATKVMIGENVDEKHLPTLDHPILPADYVAIKAPMFSWPRLRD
ADPILLRCEMASTGEVACFGEIHTAFLKAMLSFGIKPKQKILIGIQSFRPRFLGVA
EQLHNEGFKLFATEATSDWLNANNVPATPVANPVSQEQGNPSLSSIRKLIRDSIDLVI
NLPNNNTKFXHDNVYIRRTAVDSGIPLLTNFQVTKLFAEAVQKSRKVDKSLFHYQY
SAGKAA"

BASE COUNT 38448 a 20708 c 22753 g 40050 t
ORIGIN

Query Match 100.0%; Score 81; DB 9; Length 121959;

Best Local Similarity 100.0%; Pred. No. 1.4e-16;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAGCGCCACACAAATGTCCCTGCCACCCAGTGGCATGGCGGTCT 60
|||||
Db 119125 GCCACATCAGACTGGCTCAGCGCCACACAAATGTCCCTGCCACCCAGTGGCATGGCGGTCT 119184
QY 61 CAAGAAGGACAGAAATCCCAGC 81
|||||
Db 119185 CAAGAAGGACAGAAATCCCAGC 119205
RESULT 11
AC008172/c
LOCUS AC008172 183696 bp DNA linear PRI 08-NOV-2000
DEFINITION Homo sapiens BAC clone RP11-349c4 from 2, complete sequence.
AC008172
VERSION AC008172.2 GI:10140832
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 183696)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 183696)
AUTHORS Le,T.P., Threideh,J. and Strommatt,C.
TITLE The sequence of Homo sapiens BAC clone RP11-349G4
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 183696)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 183696)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 183696)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Sep 15, 2000 this sequence version replaced gi:5630064.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics

Center project name: H_NH0349G04
NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

variation 1631 /gene="CPS1" /note="mutation in CPS1 deficiency" /replace="t" 2679 /gene="CPS1" /note="putative polymorphism (silent)" /replace="g" 2838 /gene="CPS1" /note="ambiguous" /replace="c" 1266 a 967 c 1113 g 1157 t

BASE COUNT 1266 a 967 c 1113 g 1157 t

ORIGIN

Query Match 100.0%; Score 81; DB 9; Length 4503; Best Local Similarity 100.0%; Pred. No. 2.3e-16; Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
|||||

Db 4177 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 4236

QY 61 CAAGAAGGACAGAATCCGAGC 81
|||||

Db 4237 CAAGAAGGACAGAATCCGAGC 4257

RESULT 8

AR184417

LOCUS AR184417 5761 bp DNA linear PAT 20-APR-2002

DEFINITION Sequence 3 from patent US 6346382.

ACCESSION AR184417

VERSION AR184417.1 GI:20230382

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 5761)

AUTHORS Summar,M.L. and Christman,B.W.

TITLE Human carbamyl phosphate synthetase I polymorphism and diagnostic methods related thereto

JOURNAL Patent: US 6346382-A 3 12-FEB-2002;

FEATURES Location/Qualifiers

source 1. .5761 /organism="unknown"

BASE COUNT 1624 a 1210 c 1309 g 1618 t

ORIGIN

Query Match 100.0%; Score 81; DB 6; Length 5761; Best Local Similarity 100.0%; Pred. No. 2.2e-16; Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
|||||

Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 4359

QY 61 CAAGAAGGACAGAATCCGAGC 81
|||||

Db 4360 CAAGAAGGACAGAATCCGAGC 4380

RESULT 9

AF154830

LOCUS AF154830 5761 bp mRNA linear PRI 09-JUN-1999

DEFINITION Homo sapiens carbamyl phosphate synthetase I mRNA, complete cds.

ACCESSION AF154830

VERSION AF154830.1 GI:5020419

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 5761)

AUTHORS Summar,M.

TITLE Direct Submission

JOURNAL Submitted (27-MAY-1999) Pediatrics, Vanderbilt University Medical Center, DD2205 Medical Center North, Nashville, TN 37232, USA

FEATURES Location/Qualifiers

source 1. .5761 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="2" /map="2q35" 124. .4626 /note="CPS1" /codon_start=1 /product="carbamyl phosphate synthetase I" /protein_id="AAD38072.1" /db_xref="GI:5020420" /translation="MTRILTAFKVVRTLKTGFGTNTVTAHOKWKFSPRGIRLLSVKAO
TAHIVLEDGTMMKGYSGFHPSSVAGEVVFNTGLGGYPEAITDPAYKQGIILTMANPIIG
NGGAPDTALDELGLSKYLESGIKVSGLLVLVDYSKDYNHMLATKSLGWLQEEKVPA
IYGVDTRMILTIIIRDKGTMLGIEFEGQPVDFDPKNQLIAEVSTKDVKYVKGKNT
KVAVDCGKKNVIRLLVKRGAEVHLVPWNHDTKMEYDGLIAGGPNPALAEPLIQ
NVRKILESDRKEPLFGISTNLITGLAAGAKYKMSMANRGQNPVLNITNKQAFITA
QNHGYALDNTLPAGWKPLFVNVDQTNNEGIMHESKPFPAVQHPHPTPGIDTLEYLFD
SFFSLIKKGKATTITSVLPKPAVASKRVEVSKVILGSGGLSIGQAGEFDYSGSQAVK
AMKEENVKTVLMNPNIASVQTNVEVLKQADTVYFLPITPQFVTEVIKAEQPDGLILGM
GGQTALNCGVELFKRGLVKEYGVKLTGSVESIMATEDROLFSDKLNEINEKIAPSFA
VESIEDALKAADTIGYPMIRSAYALGGLSGGICPNRETMDLSTKAFAMTNQIILVEK
SVTGWKEIEYEVVRDADDNCTVCNMENVDAWGVTGDSVVVAPAQITLSNAEFQMLRR
TSINVVRHLGIVGEONIQFALHPTSMEYCIIEVNARLSRSLASLASKATGYPLAFIAAK
IALGIVPLPEIKNVVSGKTSACFEPISLDYMTVKIPRWDLDRFHGTSSRIGSSMKSVGEV
MAIGRTFEESFQKALRMCHPSIEGTPRLPMNKWPSNLDLRKELSEPSSTRIYAIK
AIDDNMSLDEIEKLTAYIDKWFLYKMRDILNMEKTLKGLNSESMTETLKRKEIGFSD
KQISKCLGLTEAQTRELRLKNIHPWKQIDTLAAEYPSVTNYLYTYNGQEHHDVNF
DHGMVILGCGPYHIGSSVEFDCAVSSIRTLRQIGKKTVVVNCNPETVSTDFDECDKL
YFEELSLEIRILDYHGEACGGCIISVGGQIPNNLAVPLYKNGVKIMGTSPLQIDRAED
RSIFSVIDELKVAQAPKAVNTLNEALEFAKSDYDPCLLRPSYVLSGSAMNVVFS
EMKKFLEATRVSQHPVLTIKFVBGAREVMDAVGDKGRVISHAISEHVEDAGVHSG
DATLMLPTQTSIQGAIEKVKDARKIAKAFASGPFNVQFLVKGNDLVIECNLRASR
SFPFVSKTLGVDFIDVATKVMIGENVDEKHLPTLDHPIIPADYVAIKAPMFSWPRLRD
ADPILRCEMASTGEVACGEGIHFAFLKAMLSTGFKIPQKGLIGIQQSFPRFLGVA
EQLHNEGFKLFATEATSDNLNANNVPATPVAPWSQEQGNPSLSIRKLIRGSDIDLVI
NLPNNNTKRVHDNVIIRRTAVDSGIPLLTNFQVTKLFAEAVQSKRKVDSKSLFHYRQY
SAGKAA"

BASE COUNT 1624 a 1210 c 1309 g 1618 t

ORIGIN

Query Match 100.0%; Score 81; DB 9; Length 5761; Best Local Similarity 100.0%; Pred. No. 2.2e-16; Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
|||||

Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCAGTGGCATGGCCGTCT 4359

QY 61 CAAGAAGGACAGAATCCGAGC 81
|||||

Db 4360 CAAGAAGGACAGAATCCGAGC 4380

RESULT 10

AF536523

LOCUS AF536523 121959 bp DNA linear PRI 02-DEC-2002

DEFINITION Homo sapiens carbamylphosphate synthetase 1 gene, complete cds; nuclear gene for mitochondrial product.

ACCESSION AF536523

VERSION AF536523.1 GI:25992605

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W. and Mannion,J.
Compositions and methods for the therapy and diagnosis of lung cancer
Patent: WO 0100828-A 93 04-JAN-2001;
CORIXA CORPORATION (US)
Location/Qualifiers
1. .531
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 157 a 139 c 108 g 127 t
ORIGIN
Query Match 100.0%; Score 81; DB 6; Length 531;
Best Local Similarity 100.0%; Pred. No. 3.1e-16;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCCGATGGCGGTCT 60
|||||
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCCGATGGCGGTCT 216
QY 61 CAAGAAGGACAGAATCCCGC 81
|||||
Db 217 CAAGAAGGACAGAATCCCGC 237
RESULT 6
AX367383
LOCUS AX367383 531 bp DNA linear PAT 17-FEB-2002
DEFINITION Sequence 93 from Patent WO0204514.
ACCESSION AX367383
VERSION AX367383.1 GI:18855487
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W., Marnerakis,M., Carter,D., Fanger,G.R., Vedvick,T.S., Bangur,C.S., McNabb,A., Fanger,N., Switzer,A., Mcneill,P.D. and Clapper,J.D.
Compositions and methods for the therapy and diagnosis of lung cancer
Patent: WO 0204514-A 93 17-JAN-2002;
CORIXA CORPORATION (US)
Location/Qualifiers
1. .531
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 157 a 139 c 108 g 127 t
ORIGIN
Query Match 100.0%; Score 81; DB 6; Length 531;
Best Local Similarity 100.0%; Pred. No. 3.1e-16;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCCGATGGCGGTCT 60
|||||
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCACCCCGATGGCGGTCT 216
QY 61 CAAGAAGGACAGAATCCCGC 81
|||||
Db 217 CAAGAAGGACAGAATCCCGC 237
RESULT 7
HSCPSTIMUT
LOCUS HSCPSTIMUT 4503 bp mRNA linear PRI 24-SEP-1998

DEFINITION Homo sapiens mRNA for carbamyl phosphate synthetase I.
ACCESSION Y15793
VERSION Y15793.1 GI:3228247
KEYWORDS carbamoyl-phosphate synthetase I; CPS1 gene; mutation.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Finckh,U., Kohlschutter,A., Schafer,H., Sperhake,K., Colombo,J.P. and Gal,A.
Prenatal diagnosis of carbamoyl phosphate synthetase I deficiency by identification of a missense mutation in CPS1
Hum. Mutat. 12 (3), 206-211 (1998)
98375696
9711878
REFERENCE 2
Finckh,U.
Direct Submission
TITLE Direct Submission
AUTHORS Submitted (05-DEC-1997) U. Finckh, Dept. of Human Genetics, University Hospital Eppendorf, Butenfeld 42, 22529 Hamburg, FRG
REMARK 3 (bases 1 to 4503)
REFERENCE
AUTHORS Finckh,U.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1998) U. Finckh, Dept. of Human Genetics, University Hospital Eppendorf, Butenfeld 42, 22529 Hamburg, FRG
On Jun 16, 1998 this sequence version replaced gi:2959899.
COMMENT Related sequence D90282.
FEATURES
Location/Qualifiers
1. .4503
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="q35"
/tissue_type="liver"
2
1. .4503
/gene="CPS1"
1. .4503
/gene="CPS1"
/codon_start=1
/product="carbamoyl phosphate synthetase 1"
/protein_id="CAA75785.1"
/db_xref="GI:3228248"
/translation="MTRILTAFKVVRILTKTGFNTNVTAHQKWFSRPGIRLLSVKQAQTAHIVLEDTKMGYSFGHPSSVAGEVFNLTGLGYPEAITDPAYKQILTWMANPIIGNGAPDTTALDELGLSKYLESNIGIXVSGLLVLDYSKDNHNLAYKSLQWLQEEKVPAIYGVDTRMLTKIIRDKGTMLGKIEFGQPVDFDPKNQNLIAEVSTKDVKYVKGKNTKVVAVDCGINKNNVIRLLVKRGAEVHLVFNHDFTKMEYDGLIAGPGNPALAEPLIQNVRKILESDRKEPLFGISTGNLITGLAAGAKTYKMSMANRGONQPVNITNKQAFITAHQGYALDNTLPAGWKPLFVNVDQNEGIMHESKPPFAVQFHPVTPGPIDEYLFDSFFSLIKKGKATTITSVLPKPAIVASRVEVSKVLLILSGGLSICQAGEFDYSGSQAAMKEENVKTVLMNPNIASVQTNEVGLKQADTVYFLPITPQFVTEVIKAEQPDGLILGMGGTALNCGVELFKRGVLKEYGVKVLGTSVESIMATEDRQLFSKLNINEKIAPSFAVESIEDALKAADTIGYPMIRSAVALGGLSGGICPNRETLMDLSTKAFAMTNOILVEKSVTGWKEIEYEVVRDADDNCVTVCNNENVNDAMGVHTGDSVVVAPAQTLNSAEFQMLRRTSINVVRHLGIVGECNIQFALHPTSMEYCIIEVNARLSRSSALASKATGYPLAFIAAKIALGIPLEIKNVVSGKTSACFEPSLDYMVTIKPRWDLDRFHGTSRIGSSMKSVGEVMAIGRTFEESFQKALRMCHPSIEGFTPLPMNKWEPNSNLDLRKELSEPSSTRIYAIKATDDNMSLDEIEKLTIDKWFLYKMRDILNMEKTLKGLNSESMTEETLKRKKEIGFSDKQISKCLGTEAQTRELRLKKNHFWVKQIDTLAAEYPSVTNLYLYTYNGQEHVDNFDHGMVVLGCGPYHIGSVVEFDWCAVSSIRTLROLGKKTVVVNCNPETVSTDFDECDKLYFEELSRLIDYHQEACGGCIIISVGGQIPNNLAVPLYKNGVKIMGTSPQLIDRAEDRSIFSAVLDELKVAQAPKAVNTLNEALEFAKSVDPYCLLRPSYVLSGSAMNVFSEDEMKKFLEBTRVTSQEHPPVLTKFVEGAREVEMDAVGKDGRLVSHAISEHVEDAGVHSGDATLMLPTQITISQGAIEKVKDATRKIAKAFASGFENVQFLVKGNDVLVIECNLRASRSPFFVSKTLGVDFIDVATKVMIGENVDEKHLPTLDHPILPADYVAIKAPMFSWPRLRDADPILRCEMASTGEVACFEGEIHATFLKAMLSTGKIPQKGLIGIQQSFRPRFLGVAEQLHNEGFLFATEATSDWLNANNVPATPVAWPSQGNPSSLIRKLIRDGSDILVINLPNNNTKPVHDNYVIRRTAVDSGIPLLTNFQVKLFAEAVQSKRKVDSKSLFHYRQYSAGKAA"
gene
CDS

TITLE Human carbamoylphosphate synthetase (CPS I) gene organization
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 296)
AUTHORS Funghini,S., Morrone,A. and Zammarchi,E.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-2002) Dep. of Paediatrics, University of Florence, Via Luca Giordano 13, Florence 50132, Italy
FEATURES
source
1. .296
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q35"
84. .196
/gene="CPSI"
/number=36
exon
BASE COUNT 93 a 64 c 48 g 91 t
ORIGIN
Query Match 100.0%; Score 81; DB 9; Length 296;
Best Local Similarity 100.0%; Pred. No. 3.4e-16;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCCGATGGCGGTCT 60
|||||
Db 99 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCCGATGGCGGTCT 158
QY 61 CAAGAAGGACAGATCCCAGC 81
|||||
Db 159 CAAGAAGGACAGATCCCAGC 179
RESULT 2
AR184418
LOCUS AR184418 495 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 5 from patent US 6346382.
ACCESSION AR184418
VERSION AR184418.1 GI:20230383
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 495)
AUTHORS Summar,M.L. and Christman,B.W.
TITLE Human carbamyl phosphate synthetase I polymorphism and diagnostic methods related thereto
JOURNAL Patent: US 6346382-A 5 12-FEB-2002;
FEATURES Location/Qualifiers
source
1. .495
/organism="unknown"
BASE COUNT 158 a 96 c 85 g 155 t 1 others
ORIGIN
Query Match 100.0%; Score 81; DB 6; Length 495;
Best Local Similarity 100.0%; Pred. No. 3.1e-16;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCCGATGGCGGTCT 60
|||||
Db 126 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCCGATGGCGGTCT 185
QY 61 CAAGAAGGACAGATCCCAGC 81
|||||
Db 186 CAAGAAGGACAGATCCCAGC 206
RESULT 3
AR272381
LOCUS AR272381 531 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 93 from patent US 6504010.
ACCESSION AR272381
VERSION AR272381.1 GI:29704266

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 531)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W., Mannion,J. and Fan,L.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: US 6504010-A 93 07-JAN-2003;
FEATURES Location/Qualifiers
source
1. .531
/organism="unknown"
BASE COUNT 157 a 139 c 108 g 127 t
ORIGIN
Query Match 100.0%; Score 81; DB 6; Length 531;
Best Local Similarity 100.0%; Pred. No. 3.1e-16;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCCGATGGCGGTCT 60
|||||
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCCGATGGCGGTCT 216
QY 61 CAAGAAGGACAGATCCCAGC 81
|||||
Db 217 CAAGAAGGACAGATCCCAGC 237
RESULT 4
AR275962
LOCUS AR275962 531 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 93 from patent US 6509448.
ACCESSION AR275962
VERSION AR275962.1 GI:29709607
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 531)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W., Mannion,J., Fan,L. and Wang,A.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: US 6509448-A 93 21-JAN-2003;
FEATURES Location/Qualifiers
source
1. .531
/organism="unknown"
BASE COUNT 157 a 139 c 108 g 127 t
ORIGIN
Query Match 100.0%; Score 81; DB 6; Length 531;
Best Local Similarity 100.0%; Pred. No. 3.1e-16;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCCGATGGCGGTCT 60
|||||
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCCGATGGCGGTCT 216
QY 61 CAAGAAGGACAGATCCCAGC 81
|||||
Db 217 CAAGAAGGACAGATCCCAGC 237
RESULT 5
AX062466
LOCUS AX062466 531 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 93 from Patent WO0100828.
ACCESSION AX062466
VERSION AX062466.1 GI:12540341
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 12, 2003, 22:25:46 ; Search time 1416.5 Seconds
(without alignments)
2339.343 Million cell updates/sec

Title: US-09-585-077C-3_COPY_4300_4380
Perfect score: 81
Sequence: 1 gccacatcagactggctcaa.....aagaaggacagaatcccagc 81

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	81	100.0	296	9	AY166970S36	AY167005 Homo sapi
2	81	100.0	495	6	AR184418	AR184418 Sequence
3	81	100.0	531	6	AR272381	AR272381 Sequence
4	81	100.0	531	6	AR275962	AR275962 Sequence
5	81	100.0	531	6	AX062466	AX062466 Sequence
6	81	100.0	531	6	AX367383	AX367383 Sequence
7	81	100.0	4503	9	HSCPSIMUT	Y15793 Homo sapien
8	81	100.0	5761	6	AR184417	AR184417 Sequence
9	81	100.0	5761	9	AF154830	AF154830 Homo sapi
10	81	100.0	121959	9	AF536523	AF536523 Homo sapi
11	81	100.0	183696	9	AC008172	AC008172 Homo sapi
12	79.4	98.0	5215	6	AX409122	AX409122 Sequence
13	79.4	98.0	5215	6	E04065	E04065 cDNA encodi
14	79.4	98.0	5215	9	HUMCPST	D90282 Homo sapien
15	79.4	98.0	5761	6	AR184416	AR184416 Sequence
16	79.4	98.0	5761	6	AR184424	AR184424 Sequence
17	79.4	98.0	5762	6	AR184425	AR184425 Sequence
18	74.6	92.1	133	10	RNCPSTIX36	227487 R.norvegicu
19	74.6	92.1	147	10	RATCPSTI11	M12327 Rattus norv
20	74.6	92.1	229369	2	AC106119	AC106119 Rattus no
21	71.4	88.1	255581	2	AC101854	AC101854 Mus muscu
22	67.4	83.2	816	6	BD091111	BD091111 DNA chip
23	48.6	60.0	4716	5	RCU05193	U05193 Rana catesb
24	47.6	58.8	580	6	AX525511	AX525511 Sequence
25	42.6	52.6	4597	5	OMU65893	U65893 Oncorhyncus
26	40	49.4	4717	5	AF006491	AF006491 Micropter
27	34.6	42.7	63058	2	AC133885	AC133885 Homo sapi
28	31.6	39.0	6001	5	SQUCARPSYN	L31362 Squalus aca
29	30.6	37.8	31	6	AX249689	AX249689 Sequence
30	29.6	36.5	4755	5	AF119250	AF119250 Oreochrom
31	29.6	36.5	63058	2	AC133885	AC133885 Homo sapi
32	29.6	36.5	72620	2	AC121127	AC121127 Mus muscu
33	29.6	36.5	161065	9	AC124799	AC124799 Homo sapi
34	29.6	36.5	161190	9	AC005137	AC005137 Human Chr
35	29.4	36.3	5989	5	AF169248	AF169248 Opsanus b
36	29.4	36.3	36307	9	HS3666D1	Z97986 Human DNA s
37	29.4	36.3	256073	9	AE006464	AE006464 Homo sapi
38	29.2	36.0	232707	2	BX323828	BX323828 Danio rer
39	29	35.8	1101	11	G67204	G67204 Xq4133 KWOK
40	29	35.8	141672	9	HS142F18	AL031073 Human DNA
41	28.8	35.6	142742	9	AF121782	AF121782 Homo sapi
42	28.8	35.6	199665	9	AF064857	AF064857 Homo sapi
43	28.8	35.6	340000	9	HS21C081	AL163281 Homo sapi
44	28.6	35.3	19795	9	HSIGCMUDE	X57331 Human immun
45	28.6	35.3	169802	9	CNS01DT2	AL122127 Human chr

ALIGNMENTS

RESULT 1
AY166970S36
LOCUS AY166970S36 296 bp DNA linear PRI 28-FEB-2003
DEFINITION Homo sapiens carbanoylphosphate synthetase (CPSI) gene, exon 36;
nuclear gene for mitochondrial product.
ACCESSION AY167005
VERSION AY167005.1 GI:28625483
KEYWORDS
SEGMENT 36 of 38
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 296)
AUTHORS Funghini,S., Morrone,A. and zammarchi,E.

```
; PRIOR APPLICATION NUMBER: US 09/629,616
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 09/494,359
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: JP 11-24149
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1113
; TYPE: PRF
; ORGANISM: Brevibacterium lactofermentum
US-09-836-470B-3

Query Match      21.3%; Score 58; DB 9; Length 1113;
Best Local Similarity 31.4%; Pred. No. 1e+02;
Matches 16; Conservative 8; Mismatches 27; Indels 0; Gaps 0;

QY      1 QLNHEGFKLFATEATSDWLNANNVPANVPVAPWPSQEQNPSSLSSIRKLIRDG 51
       :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      994 RLASMGYKILATEGTAGMLRRNGIDCEVVLKASDIREGVEGKSIYDRIREG 1044

RESULT 14
US-10-284-138-3
; Sequence 3, Application US/10284138
; Publication No. US20030082774A1
; GENERAL INFORMATION:
; APPLICANT: KUWABARA, Yoko
; APPLICANT: HASHIGUCHI, Kenichi
; APPLICANT: NAKAMATSU, Tsuyoshi
; APPLICANT: KURAHASHI, Osamu
; APPLICANT: MORI, Yukiko
; APPLICANT: ITO, Hisao
; TITLE OF INVENTION: CARBAMOYL-PHOSPHATE SYNTHETASE GENE OF CORYNEFORM
; TITLE OF INVENTION: BACTERIA AND METHOD FOR
; TITLE OF INVENTION: PRODUCING L-ARGININE
; FILE REFERENCE: 0010-1082-0
; CURRENT APPLICATION NUMBER: US/10/284,138
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US/09/494,359
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: JP 11-24149
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1113
; TYPE: PRF
; ORGANISM: Brevibacterium lactofermentum
US-10-284-138-3

Query Match      21.3%; Score 58; DB 15; Length 1113;
Best Local Similarity 31.4%; Pred. No. 1e+02;
Matches 16; Conservative 8; Mismatches 27; Indels 0; Gaps 0;

QY      1 QLNHEGFKLFATEATSDWLNANNVPANVPVAPWPSQEQNPSSLSSIRKLIRDG 51
       :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      994 RLASMGYKILATEGTAGMLRRNGIDCEVVLKASDIREGVEGKSIYDRIREG 1044

RESULT 15
US-10-284-334-3
; Sequence 3, Application US/10284334
; Publication No. US20030082775A1
; GENERAL INFORMATION:
; APPLICANT: KUWABARA, Yoko
; APPLICANT: HASHIGUCHI, Kenichi
; APPLICANT: NAKAMATSU, Tsuyoshi
; APPLICANT: KURAHASHI, Osamu
; APPLICANT: MORI, Yukiko
; APPLICANT: ITO, Hisao
; TITLE OF INVENTION: CARBAMOYL-PHOSPHATE SYNTHETASE GENE OF CORYNEFORM
```

```
; TITLE OF INVENTION: BACTERIA AND METHOD FOR
; TITLE OF INVENTION: PRODUCING L-ARGININE
; FILE REFERENCE: 0010-1082-0
; CURRENT APPLICATION NUMBER: US/10/284,334
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US/09/494,359
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: JP 11-24149
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1113
; TYPE: PRF
; ORGANISM: Brevibacterium lactofermentum
US-10-284-334-3

Query Match      21.3%; Score 58; DB 15; Length 1113;
Best Local Similarity 31.4%; Pred. No. 1e+02;
Matches 16; Conservative 8; Mismatches 27; Indels 0; Gaps 0;

QY      1 QLNHEGFKLFATEATSDWLNANNVPANVPVAPWPSQEQNPSSLSSIRKLIRDG 51
       :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      994 RLASMGYKILATEGTAGMLRRNGIDCEVVLKASDIREGVEGKSIYDRIREG 1044

Search completed: September 13, 2003, 02:28:03
Job time : 39 secs
```

RESULT 11
US-10-169-048-40
; Sequence 40, Application US/10169048
; Publication NO. US20030072769A1
; GENERAL INFORMATION:
; APPLICANT: Clarke, Edna Elizabeth
; APPLICANT: Zhou, Liqing
; APPLICANT: Shea, Jacqueline Elizabeth
; APPLICANT: Feldman, Robert Graham
; APPLICANT: Holden, David William
; TITLE OF INVENTION: Streptococcus P

RESULT 13
US-09-836-470B-3
; Sequence 3, Application US/09836470B
; Patent No. US20020090702A1
; GENERAL INFORMATION:
; APPLICANT: KOWABARA, Yoko
; APPLICANT: HASHIGUCHI, , kenichi
; APPLICANT: NAKAMATSU, Tsuyoshi
; APPLICANT: KURAHASHI, Osamu
; APPLICANT: MORI, Yukiko
; APPLICANT: ITO, Hisao
; TITLE OF INVENTION: CARBAMOYL-PHOSP
; TITLE OF INVENTION: PRODUCING L-AR
; FILE REFERENCE: 205801US0DIV
; CURRENT APPLICATION NUMBER: US/09/8
; CURRENT FILING DATE: 2001-04-18

;
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474

;
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14

; FEATURE:
; OTHER INFORMATION: Xaa = Unknown or Other at position 174
; OTHER INFORMATION: Xaa = Unknown or Other at position 363
; OTHER INFORMATION: Xaa = Unknown or Other at position 1088
; OTHER INFORMATION: Xaa = Unknown or Other at position 1131
; OTHER INFORMATION: Xaa = Unknown or Other at position 1217
; OTHER INFORMATION: Xaa = Unknown or Other at position 1389
US-09-468-147-91

Query Match 22.2%; Score 60.5; DB 11; Length 1698;
Best Local Similarity 44.4%; Pred. No. 81;
Matches 12; Conservative 4; Mismatches 10; Indels 1; Gaps 1;

QY 3 HNEGFKLFATEATSDWLNANNVPANPV 29
| | :| :| | | | | | | | | |
Db 911 HRPGEDELYLTPEAANWFEANK-PAQPV 936

RESULT 6
US-09-764-878-129
; Sequence 129, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL21
; CURRENT APPLICATION NUMBER: US/09/764,878
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 129
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-878-129

Query Match 22.1%; Score 60; DB 9; Length 70;
Best Local Similarity 52.4%; Pred. No. 2.2;
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 31 WPSQEQGNPSLSIRKLIRDG 51
:| | | :| | | | :| | :| |
Db 39 YPGQHGETPSSLKIQKLVRCG 59

RESULT 7
US-09-764-860-464
; Sequence 464, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 464
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-860-464

Query Match 22.1%; Score 60; DB 9; Length 70;
Best Local Similarity 52.4%; Pred. No. 2.2;
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 31 WPSQEQGNPSLSIRKLIRDG 51
:| | | :| | | | :| | :| |
Db 39 YPGQHGETPSSLKIQKLVRCG 59

RESULT 8
US-10-079-854-129
; Sequence 129, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL21C1
; CURRENT APPLICATION NUMBER: US/10/079,854
; CURRENT FILING DATE: 2002-02-22
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 129
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-079-854-129

Query Match 22.1%; Score 60; DB 15; Length 70;
Best Local Similarity 52.4%; Pred. No. 2.2;
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 31 WPSQEQGNPSLSIRKLIRDG 51
:| | | :| | | | :| | :| |
Db 39 YPGQHGETPSSLKIQKLVRCG 59

RESULT 9
US-10-074-095-464
; Sequence 464, Application US/10074095
; Publication No. US20030077704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: 09/764,860
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08

US-09-815-242-5059

Query Match	24.8%;	Score 67.5;	DB 9;	Length 344;
Best Local Similarity	37.1%;	Pred. No. 1.6;		
Matches 13;	Conservative	7;	Mismatches 14;	Indels 1;
				Gaps 1;

QY	1 QLNHEGFKLFATEATSDWL-NANNVPANPVAVPSQ 34 ::: : :::
Db	61 RVHSEGFVRFLQNAWODLWLTGRSHDMLPIAWETR 95

```

RESULT 2
US-10-156-761-14395
; Sequence 14395, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14395
; LENGTH: 1102
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14395

```

Query Match	24.3%;	Score 66;	DB 15;	Length 1102;
Best Local Similarity	38.5%;	Pred. No. 9.6;		
Matches 20;	Conservative	8;	Mismatches 22;	Indels 2;
			Gaps	2;

QY	1	QLHNEGFKLFATEATSDMLNANNVPANVPWAPWSQEQNP-SLSIRKLLIRDG	51
		: : : :	
DQ	981	ELVAHGFELLATSGTAEYLKRNGINATVVRKQS-EGEGPGGKTVQLIHDG	1031

```

RESULT 3
US-09-976-059-30
; Sequence 30, Application US/09976059
; Patent No. US20020164747A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Zazopoulos, Emmanuel
; APPLICANT: Staffa, Alfredo
; TITLE OF INVENTION: Genes and Proteins for Biosynthesis of Ramoplanin
; FILE REFERENCE: 3019-PCT
; CURRENT APPLICATION NUMBER: US/09/976,059
; CURRENT FILING DATE: 2001-10-15
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Actinoplanes sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: V represents a non-standard initiator codon. It
; OTHER INFORMATION: the biosynthesized protein will have a formylmeth
; OTHER INFORMATION: at this position
US-09-976-059-30

```

Query Match	23.2%	Score 63;	DB 10;	Length 619;
Best Local Similarity	34.5%	Pred. No. 12;		
Matches 19:	Conservative	7;	Mismatches 21;	Indels 8;
				Gaps 3;

3 HNEGFKLFATEATSDWLNANN--VPAN---PVAWPSEQGNPSLSIRKLLIRDG 51
| : | ||| | | | : | ||| :|||
514 HRDGRTYLA--ATSSNRTAGRLIVPTGHSYPLGGFGSAAFPFSLAGVQRLVRDG 566

```

RESULT 4
US-09-117-447-2
; Sequence 2, Application US/09117447
; Patent No. US20020168728A1
; GENERAL INFORMATION:
; APPLICANT: LUBITZ, Werner
; APPLICANT: SLEYTR, Uwe
; APPLICANT: KUEN, Beatrix
; APPLICANT: TRUPPE, Michaela
; APPLICANT: HOWORKA, Stefan
; APPLICANT: RESCH, Stepanka
; APPLICANT: SCHROLL, Gerhard
; APPLICANT: SARA, Margit
; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF S-LAYER PROTEINS
; FILE REFERENCE: 100564-08013
; CURRENT APPLICATION NUMBER: US/09/117,447
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: PCT/EP97/00432
; PRIOR FILING DATE: 1997-01-31
; PRIOR APPLICATION NUMBER: DE/196 03 649.6
; PRIOR FILING DATE: 1996-02-01
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1228
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus
US-09-117-447-2

```

Query Match	22.2%;	Score 60.5;	DB 10;	Length 1228;
Best Local Similarity	29.4%;	Pred. No. 55;		
Matches 15: Conservative	11;	Mismatches 20;	Indels 5;	Gaps 1;

QY 4 NEGKLE-----ATEATSDWLNANNVPANPVAMPSQEQGNPLSSIRKLIR 49
 : : : : :
Db 711 NOGYKLWFEGKATAPVKDAANANTLATNYIYTFTTEGODVTAPTVTKVFK 761

```

RESULT 5
US-09-468-147-91
; Sequence 91, Application US/09468147A
; Publication No. US20030049601A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232.US.P1
; CURRENT APPLICATION NUMBER: US/09/468,147A
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/173,141
; EARLIER FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: US 60/061,199
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 91
; LENGTH: 1698
; TYPE: PRT
; ORGANISM: Hepatitis E Virus

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2003, 02:16:53 ; Search time 36 Seconds
(without alignments)
206.709 Million cell updates/sec

Title: US-09-585-077C-2_COPY_1380_1430
 Perfect score: 272
 Sequence: 1 QLNHEGFKLFATEATSDWLN.....PSQEQNPSSIRKLIRDG 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 541936 seqs, 145912426 residues

Total number of hits satisfying chosen parameters: 541936

```

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10
                  Listing first 45

```

```
Database :
Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query ⁸			DB	ID	Description
	Score	Match	Length			
1	67.5	24.8	344	9	US-09-815-242-5059	Sequence 5059, Ap
2	66	24.3	1102	15	US-10-156-761-14395	Sequence 14395, A
3	63	23.2	619	10	US-09-976-059-30	Sequence 30, Appl
4	60.5	22.2	1228	10	US-09-117-447-2	Sequence 2, Appl
5	60.5	22.2	1698	11	US-09-468-147-91	Sequence 91, Appl
6	60	22.1	70	9	US-09-764-878-129	Sequence 129, App
7	60	22.1	70	9	US-09-764-860-464	Sequence 464, App
8	60	22.1	70	15	US-10-079-854-129	Sequence 129, App
9	60	22.1	70	15	US-10-074-095-464	Sequence 464, App
10	59.5	21.9	504	15	US-10-156-761-12417	Sequence 12417, A
11	59	21.7	339	15	US-10-169-048-40	Sequence 40, Appl
12	58	21.3	1018	11	US-09-494-359-3	Sequence 3, Appl
13	58	21.3	1113	9	US-09-836-470B-3	Sequence 3, Appl
14	58	21.3	1113	15	US-10-284-138-3	Sequence 3, Appl
15	58	21.3	1113	15	US-10-284-334-3	Sequence 3, Appl

16	57.5	21.1	1435	15	US-10-128-714-8125	Sequence 8125, Ap
17	57	21.0	100	10	US-09-872-523-8	Sequence 8, Appli
18	56.5	20.8	576	14	US-10-029-180-102	Sequence 102, App
19	56	20.6	1113	10	US-09-738-626-5279	Sequence 5279, Ap
20	55.5	20.4	162	15	US-10-156-761-13789	Sequence 13789, A
21	55.5	20.4	232	15	US-10-156-761-11422	Sequence 11422, A
22	55.5	20.4	823	15	US-10-177-293-240	Sequence 240, App
23	55	20.2	1626	12	US-10-202-167-2	Sequence 2, Appli
24	54.5	20.0	639	10	US-09-906-209-2	Sequence 2, Appli
25	54.5	20.0	673	15	US-10-157-031-291	Sequence 291, App
26	54	19.9	441	9	US-09-879-957-34	Sequence 34, Appli
27	54	19.9	570	9	US-09-815-242-10760	Sequence 10760, A
28	54	19.9	1012	10	US-09-285-385C-4	Sequence 4, Appli
29	53.5	19.7	766	15	US-10-156-761-10075	Sequence 10075, A
30	53.5	19.7	1708	11	US-09-468-147-166	Sequence 166, App
31	53	19.5	140	12	US-10-238-075-984	Sequence 984, App
32	53	19.5	274	12	US-09-882-227-196	Sequence 196, App
33	53	19.5	395	15	US-10-128-714-3592	Sequence 3592, Ap
34	53	19.5	620	15	US-10-128-714-8592	Sequence 8592, A
35	52.5	19.3	311	15	US-10-156-761-14936	Sequence 14936, A
36	52.5	19.3	346	15	US-10-102-806-577	Sequence 577, App
37	52	19.1	644	10	US-09-943-692-2	Sequence 2, Appli
38	52	19.1	644	15	US-10-222-441-2	Sequence 2, Appli
39	52	19.1	644	15	US-10-229-346-2	Sequence 2, Appli
40	52	19.1	844	14	US-10-055-364-40	Sequence 40, Appl
41	52	19.1	1601	10	US-09-862-027-40	Sequence 40, Appl
42	52	19.1	3739	10	US-09-861-289-33	Sequence 33, Appl
43	52	19.1	3739	10	US-09-860-846-33	Sequence 33, Appl
44	52	19.1	3739	11	US-09-988-384B-33	Sequence 33, Appl
45	52	19.1	3739	11	US-09-836-821-33	Sequence 33, Appl

ALIGNMENTS

```

RESULT 1
US-09-815-242-5059
; Sequence 5059, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5059
; LENGTH: 344
; TYPE: PRY
; ORGANISM: Pseudomonas aeruginosa

```



```
; SEQ ID NO 4
; LENGTH: 1012
; TYPE: PRT
; ORGANISM: mouse
US-09-285-385C-4

Query Match      19.9%; Score 54; DB 4; Length 1012;
Best Local Similarity 23.7%; Pred. No. 1.5e+02;
Matches 14; Conservative 7; Mismatches 26; Indels 12; Gaps 2;

QY      2 LHNEGFKLFATEATSDWL-----NANNVPANPVAVWPSQEQNPSSLSSIRKLIIRDG 51
      | : ||| : || : : : || : | : : ||
Db      65 LDEDDLKLFHIDKAEDWTKPSIDKPGHDTGGLETSARWPNDTASNASIQAPRK---DG 120

RESULT 15
US-09-462-606-2
; Sequence 2, Application US/09462606
; Patent No. 6432408
; GENERAL INFORMATION:
; APPLICANT: MENG, XIANG-JIN
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
; FILE REFERENCE: 20264267US1
; CURRENT APPLICATION NUMBER: US/09/462,606
; CURRENT FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 60/053069
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: PCT/US98/14665
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1708
; TYPE: PRT
; ORGANISM: Hepatitis E virus
US-09-462-606-2

Query Match      19.7%; Score 53.5; DB 4; Length 1708;
Best Local Similarity 42.3%; Pred. No. 3.4e+02;
Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY      3 HNEGFKLFATEATSDWLNANNVPANP 28
      | : | : | : | : | : | : | : |
Db      921 HRPGEDELYLTPAAAWFEANK-PAQP 945

Search completed: September 13, 2003, 02:17:33
Job time : 22 secs
```

```

; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-403-618A-8
Query Match      20.0%; Score 54.5; DB 4; Length 878;
Best Local Similarity 29.8%; Pred. No. 1.1e+02;
Matches 14; Conservative 6; Mismatches 20; Indels 7; Gaps 2;

QY 11 ATEATSDWL-NANNVPANPVAMPSEQGNPSLSSIRKL-----IRD 50
   :| | | | | | | | | | | | | | | | | | | | |
Db 210 SANSTRWLVDINNETLPLAWDGLAAEKPRLDSPFSDISIYELHIRD 256

RESULT 12
US-08-630-915A-34
; Sequence 34, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-630-915A-34
Query Match      19.9%; Score 54; DB 4; Length 441;
Best Local Similarity 37.5%; Pred. No. 53;
Matches 9; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 18 WLNANNVPANPVAMPSEQGNPSL 41
   | : | | | | | | | | | |
Db 283 WFRSTSGPGMPMNWPQFEWNPDL 306

RESULT 13
US-09-107-532A-6689
; Sequence 6689, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
```

```

; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6689:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 569 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...569
; SEQUENCE DESCRIPTION: SEQ ID NO: 6689:
US-09-107-532A-6689
Query Match      19.9%; Score 54; DB 4; Length 569;
Best Local Similarity 23.9%; Pred. No. 73;
Matches 11; Conservative 12; Mismatches 23; Indels 0; Gaps 0;

QY 2 LHNEGPKLFATEATSDWLNANNVPANPVAMPSEQGNPSLSSIRKL 47
   | | : : | | | | | | | | | | | | | | | |
Db 466 LQNDQLTMKTAKKTDYWLHAKNIPGSHVIKSDKPSDETIIEAEL 511

RESULT 14
US-09-285-385C-4
; Sequence 4, Application US/09285385C
; Patent No. 6579702
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S.
; APPLICANT: Scott, Ian C.
; APPLICANT: Thomas, Christina L.
; TITLE OF INVENTION: MAMMALIAN TOLLOID-LIKE GENE AND PROTEIN
; FILE REFERENCE: 960296.96111
; CURRENT APPLICATION NUMBER: US/09/285,385C
; CURRENT FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/111873
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/080550
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
```

```
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001c-5495

Query Match      23.2%; Score 63; DB 4; Length 228;
Best Local Similarity 26.0%; Pred. No. 1.5;
Matches 13; Conservative 11; Mismatches 22; Indels 4; Gaps 1;

Qy 1 QLNHEGFKLFATEATSDWLNANNVPANVPWPSQEQNPSSLSSIRKLIRD 50
    :| | | : | :| | : : :| | : | :| |
Db 157 KLKEEDFDILGIENVINWLTKSNIKSKDIS---EGTNVEIEKIELLLKD 202

RESULT 8
US-09-252-991A-20252
; Sequence 20252, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20252
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20252

Query Match      22.4%; Score 61; DB 4; Length 167;
Best Local Similarity 37.9%; Pred. No. 1.9;
Matches 11; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

Qy 12 TQATSDWLNANNVPANVPWPSQEQNPSS 40
    :| | : | :| | | | :| |
Db 127 SSATAPWSRPSAAPANSAGWPSSRRWPS 155

RESULT 9
US-09-252-991A-27861
; Sequence 27861, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27861
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27861

Query Match      21.3%; Score 58; DB 4; Length 96;
Best Local Similarity 40.6%; Pred. No. 2.4;
Matches 13; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

Qy 14 ATSDWLNANNVPANVPWPSQEQNPSSLSSIR 45
    :| | :| | | | | | | | | |
Db 47 SSSPWPATRPATSVAVTSPACTRPSCSYR 78
```

```
RESULT 10
US-09-629-616-3
; Sequence 3, Application US/09629616
; Patent No. 6255086
; GENERAL INFORMATION:
; APPLICANT: KUWABARA, Yoko
; APPLICANT: HASHIGUCHI, Kenichi
; APPLICANT: NAKAMATSU, Tsuyoshi
; APPLICANT: KURAHASHI, Osamu
; APPLICANT: MORI, Yukiko
; APPLICANT: ITO, Hisao
; TITLE OF INVENTION: CARBAMOYL-PHOSPHATE SYNTHETASE GENE OF CORYNEFORM
; TITLE OF INVENTION: BACTERIA AND METHOD FOR PRODUCING L-ARGININE
; FILE REFERENCE: OP945CIP
; CURRENT APPLICATION NUMBER: US/09/629,616
; CURRENT FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1113
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum
US-09-629-616-3

Query Match      21.3%; Score 58; DB 3; Length 1113;
Best Local Similarity 31.4%; Pred. No. 50;
Matches 16; Conservative 8; Mismatches 27; Indels 0; Gaps 0;

Qy 1 QLNHEGFKLFATEATSDWLNANNVPANVPWPSQEQNPSSLSSIRKLIRDG 51
    :| :| :| :| | :| :| :| :| :| :| :| :|
Db 994 RLASMGYKILATEGAGMLRRNGIDCEVVLKASDIREGVEGKSIVDRIREG 1044

RESULT 11
US-09-403-618A-8
; Sequence 8, Application US/09403618A
; Patent No. 6429358
; GENERAL INFORMATION:
; APPLICANT: BROGLIE, KAREN E.
; TITLE OF INVENTION: CORN PULLULANASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95 (7.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/403,618A
; FILING DATE: 05-No. 6429358-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/045,723
; FILING DATE: MAY 6, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MAJARIAN, WILLIAM R.
; REGISTRATION NUMBER: 41,173
; REFERENCE/DOCKET NUMBER: BB-1108
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 878 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2003, 02:12:52 ; Search time 19 Seconds
(without alignments)
113.571 Million cell updates/sec

Title: US-09-585-077C-2_COPY_1380_1430
Perfect score: 272
Sequence: 1 QLNHEGFKLFATEATSDWLN.....PSQEQNPSSLSSIRKLIRDG 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA: *
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	272	100.0	1500	4	US-09-323-472A-2
2	272	100.0	1500	4	US-09-323-472A-12
3	272	100.0	1500	4	US-09-323-472A-14
4	266	97.8	1500	4	US-09-323-472A-4
5	67.5	24.8	349	4	US-09-252-991A-27544
6	63	23.2	228	4	US-09-134-001C-4694
7	63	23.2	228	4	US-09-134-001C-5495
8	61	22.4	167	4	US-09-252-991A-20252
9	58	21.3	96	4	US-09-252-991A-27861
10	58	21.3	1113	3	US-09-629-616-3
11	54.5	20.0	878	4	US-09-403-618A-8
12	54	19.9	441	4	US-08-630-915A-34
13	54	19.9	569	4	US-09-107-532A-6689
14	54	19.9	1012	4	US-09-285-385C-4
15	53.5	19.7	1708	4	US-09-462-606-2
16	53	19.5	141	4	US-09-634-238-290
17	53	19.5	246	4	US-09-252-991A-25947
18	52.5	19.3	329	4	US-09-252-991A-30436
19	52	19.1	507	4	US-09-130-337A-25
20	52	19.1	644	1	US-07-628-788A-4
21	52	19.1	644	1	US-08-072-281-2
22	52	19.1	644	1	US-08-295-060-2
23	52	19.1	644	1	US-08-759-446-2
24	52	19.1	644	3	US-09-027-998A-2
25	52	19.1	644	4	US-09-696-348-2
26	52	19.1	644	5	PCT-US92-11337-4
27	52	19.1	645	1	US-08-295-060-4

28	52	19.1	652	3	US-08-996-441B-113	Sequence 113, App
29	52	19.1	652	3	US-08-993-722A-113	Sequence 113, App
30	52	19.1	652	3	US-08-993-170A-113	Sequence 113, App
31	52	19.1	652	3	US-08-993-775B-113	Sequence 113, App
32	52	19.1	874	3	US-08-804-439A-15	Sequence 15, Appl
33	52	19.1	874	3	US-08-720-229-15	Sequence 15, Appl
34	52	19.1	1037	4	US-09-134-001C-4794	Sequence 4794, Ap
35	52	19.1	1601	4	US-09-345-473E-40	Sequence 40, Appl
36	52	19.1	3739	3	US-09-320-878-2	Sequence 2, Appl
37	52	19.1	3739	3	US-09-105-537-33	Sequence 33, Appl
38	52	19.1	3739	4	US-09-141-908-3	Sequence 3, Appl
39	52	19.1	3739	4	US-09-657-440-2	Sequence 2, Appl
40	52	19.1	11877	3	US-09-105-537-6	Sequence 6, Appl
41	51.5	18.9	303	4	US-09-252-991A-22757	Sequence 22757, A
42	51.5	18.9	859	4	US-09-107-532A-6829	Sequence 6829, Ap
43	51	18.8	196	4	US-09-252-991A-16779	Sequence 16779, A
44	51	18.8	227	6	5498499-2	Patent No. 5498499
45	51	18.8	354	4	US-09-198-452A-536	Sequence 536, App

ALIGNMENTS

RESULT 1
US-09-323-472A-2
; Sequence 2, Application US/09323472A
; Patent No. 6346382
; GENERAL INFORMATION:
; APPLICANT: Summar, Marshall
; APPLICANT: Christman, Brian
; TITLE OF INVENTION: HUMAN CARBAMYL PHOSPHATE SYNTHETASE I POLYMORPHISM AND DIAGN
; FILE REFERENCE: 1242/19
; CURRENT APPLICATION NUMBER: US/09/323,472A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1500
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-323-472A-2

Query Match 100.0%; Score 272; DB 4; Length 1500;
Best Local Similarity 100.0%; Pred. No. 5.7e-27;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLNHEGFKLFATEATSDWLNANNVPANPVAVWPSQEQNPSSLSSIRKLIRDG 51
|||||
Db 1380 QLNHEGFKLFATEATSDWLNANNVPANPVAVWPSQEQNPSSLSSIRKLIRDG 1430

RESULT 2
US-09-323-472A-12
; Sequence 12, Application US/09323472A
; Patent No. 6346382
; GENERAL INFORMATION:
; APPLICANT: Summar, Marshall
; APPLICANT: Christman, Brian
; TITLE OF INVENTION: HUMAN CARBAMYL PHOSPHATE SYNTHETASE I POLYMORPHISM AND DIAGN
; FILE REFERENCE: 1242/19
; CURRENT APPLICATION NUMBER: US/09/323,472A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 1500
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-323-472A-12
Query Match 100.0%; Score 272; DB 4; Length 1500;

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96051398; PubMed=7584044;
RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
RA Seki N., Kawarabayashi Y., Ishikawa K., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. II.
RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 1:223-229(1994).
DR EMBL; D38521; BAA07526.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 1798 AA; 206406 MW; 6FCB2F2AA77999C1 CRC64;

Query Match 23.3%; Score 63.5; DB 4; Length 1798;
Best Local Similarity 37.8%; Pred. No. 66;
Matches 14; Conservative 7; Mismatches 9; Indels 7; Gaps 2;

QY 1 QLHNEGKLFATEATSDWLNANNVPANPVAVPWSQEGQ 37
Db :||:||||: | |:: |:: | |||
180 ELHHKGFKLWFEDELIGLWVSVQNLP-----QW---EGQ 209

RESULT 15
Q9ESY1
ID Q9ESY1 PRELIMINARY; PRT; 1803 AA.
AC Q9ESY1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TEMO.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Mruk D., Mo M.-Y., Cheng C.Y.;
RT "TEMO is a marker to study sertoli-germ cell interactions: cloning and
RT regulation of a novel testicular molecule.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF296169; AAC09060.1; -.
SQ SEQUENCE 1803 AA; 206588 MW; ED73B36E135F25B2 CRC64;

Query Match 23.3%; Score 63.5; DB 11; Length 1803;
Best Local Similarity 37.8%; Pred. No. 66;
Matches 14; Conservative 7; Mismatches 9; Indels 7; Gaps 2;

QY 1 QLHNEGKLFATEATSDWLNANNVPANPVAVPWSQEGQ 37
Db :||:||||: | |:: |:: | |||
185 ELHHKGFKLWFEDELIGLWVSVQNLP-----QW---EGQ 214

Search completed: September 13, 2003, 02:15:42
Job time : 65.5 secs

RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE009081; AAL42195.1; -.
DR EMBL; AE008047; AAK86987.1; -.
DR InterPro; IPR002092; RNA_pol_phage.
DR Pfam; PF00940; RNA_pol; 1.
DR PROSITE; PS00900; RNA_POL_PHAGE_1; 1.
DR PROSITE; PS00489; RNA_POL_PHAGE_2; 1.
KW DNA-directed RNA polymerase; Complete proteome.
SQ SEQUENCE 863 AA; 97075 MW; 9B038C4FB4D36CA CRC64;

Query Match 23.5%; Score 64; DB 16; Length 863;
Best Local Similarity 30.2%; Pred. No. 24;
Matches 19; Conservative 7; Mismatches 23; Indels 14; Gaps 2;

QY 3 HNEGFKLEATEATSDWL--NANNVPANVPVWPVWGQGNPSLSSIR-----KLI 48
|| : ||: |||| | : ||| | : ||| | : ||| : |||
Db 221 HNTNIYVATKETLDWLATENSRLAPLSVYLPLVPPRPWTSFRRGGYWSGRVNLRLI 280

QY 49 RDG 51
: |
Db 281 KTG 283

RESULT 11
Q9W709
ID Q9W709 PRELIMINARY; PRT; 339 AA.
AC Q9W709;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Transposase.
OS Pleuronectes platessa (Plaice).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidei; Pleuronectidae; Pleuronectes.
OX NCBI_TaxID=8262;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=Tcl-like;
RA Leaver M.J.;
RT "A family of closely related Tcl-like transposons from diverse aquatic
RT organisms.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ249084; CAB51371.1; -.
DR InterPro; IPR002492; Transposase_5.
DR Pfam; PF01498; Transposase_5; 1.
DR ProDom; PD002059; Transposase_5; 1.
SQ SEQUENCE 339 AA; 38809 MW; 75FD9FCE9E27EFF3 CRC64;

Query Match 23.3%; Score 63.5; DB 13; Length 339;
Best Local Similarity 40.5%; Pred. No. 9.6;
Matches 15; Conservative 5; Mismatches 12; Indels 5; Gaps 2;

QY 14 ATSDWLNANNVPANVPVWPVWGQGNPSLSSIRKLIRD 50
|| : || | : |||| | : ||| | : ||| | : |||
Db 252 ATKWFGLKNV--NVLKWPS---QSPDLNPIENLWQD 283

RESULT 12
Q98TI8
ID Q98TI8 PRELIMINARY; PRT; 339 AA.
AC Q98TI8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Putative transposase.
GN TN6.
OS Pleuronectes platessa (Plaice).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidei; Pleuronectidae; Pleuronectes.
OX NCBI_TaxID=8262;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=Tcl;
RA Leaver M.J.;
RT "A family of Tcl-like transposons from the genomes of fishes and
RT frogs.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ303069; CAC28060.1; -.
DR InterPro; IPR002492; Transposase_5.
DR Pfam; PF01498; Transposase_5; 1.
DR ProDom; PD002059; Transposase_5; 1.
SQ SEQUENCE 339 AA; 38882 MW; 145224D01D64442D CRC64;

Query Match 23.3%; Score 63.5; DB 13; Length 339;
Best Local Similarity 40.5%; Pred. No. 9.6;
Matches 15; Conservative 5; Mismatches 12; Indels 5; Gaps 2;

QY 14 ATSDWLNANNVPANVPVWPVWGQGNPSLSSIRKLIRD 50
|| : || | : ||| | : ||| | : ||| | : |||
Db 252 ATKWFGLKNV--NVLKWPS---QSPDLNPIENLWQD 283

RESULT 13
Q9PW00
ID Q9PW00 PRELIMINARY; PRT; 339 AA.
AC Q9PW00;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Transposase.
OS Pleuronectes platessa (Plaice).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidei; Pleuronectidae; Pleuronectes.
OX NCBI_TaxID=8262;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=Tcl-like;
RA Leaver M.J.;
RT "A family of closely related Tcl-like transposons from diverse aquatic
RT organisms.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ249085; CAB51372.1; -.
DR InterPro; IPR002492; Transposase_5.
DR Pfam; PF01498; Transposase_5; 1.
DR ProDom; PD002059; Transposase_5; 1.
SQ SEQUENCE 339 AA; 38908 MW; 732735643B27EFF3 CRC64;

Query Match 23.3%; Score 63.5; DB 13; Length 339;
Best Local Similarity 40.5%; Pred. No. 9.6;
Matches 15; Conservative 5; Mismatches 12; Indels 5; Gaps 2;

QY 14 ATSDWLNANNVPANVPVWPVWGQGNPSLSSIRKLIRD 50
|| : || | : ||| | : ||| | : ||| | : |||
Db 252 ATKWFGLKNV--NVLKWPS---QSPDLNPIENLWQD 283

RESULT 14
Q14997
ID Q14997 PRELIMINARY; PRT; 1798 AA.
AC Q14997;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein KIAA0077 (Fragment).
GN KIAA0077.
OS Homo sapiens (Human).

```
ID O93937 PRELIMINARY; PRT; 2275 AA.
AC O93937;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE PYRABCN (EC 6.3.5.5).
GN PYRABCN.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=162425;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FGSC 4;
RA Aleksenko A., Liu W., Gojkovic Z., Nielsen J., Piskur J.;
RT "Structural and transcriptional analysis of the pyrABCN, pyrD and pyrF
RT genes in Aspergillus nidulans and the evolutionary origin of fungal
RT dihydroorotases.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF112473; AAD09129.1; -.
DR HSSP; P00968; 1CS0.
DR InterPro; IPR006220; Anth_synthII.
DR InterPro; IPR006130; Asp/Orn_Cotranf.
DR InterPro; IPR002082; Asp_carbmltransf.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; CPase_L.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005480; CPase_L_D3.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR002474; CPSase_sm_chain.
DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR000991; GATase_L.
DR InterPro; IPR004362; MGS_like.
DR InterPro; IPR006131; OTCace_O.
DR InterPro; IPR006132; OTCace_P.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 3.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR Pfam; PF02142; MGS; 1.
DR Pfam; PF00185; OTCace; 1.
DR Pfam; PF02729; OTCace_N; 1.
DR PRINTS; PR00097; ANTSNTHASEII.
DR PRINTS; PR00100; AOTCASE.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR PRINTS; PR00096; GATASE.
DR TIGRFAMS; TIGR00670; asp_carb_tr; 1.
DR TIGRFAMS; TIGR01368; CPSaseII_small; 1.
DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
DR PROSITE; PS00442; GATASE_TYPE_I; 1.
KW Ligase.
SQ SEQUENCE 2275 AA; 249739 MW; 27BA9C1FA751436A CRC64;

Query Match 25.4%; Score 69; DB 3; Length 2275;
Best Local Similarity 52.0%; Pred. No. 17;
Matches 13; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 QLNHEGFKLFATEATSDWLNANNVP 25
:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 1422 KLRDVGFEFLFATSGTADFLKENGVP 1446

RESULT 9
Q9I6H0
ID Q9I6H0 PRELIMINARY; PRT; 344 AA.
AC Q9I6H0;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
```

```
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Probable acetylpolymine aminohydrolase.
GN PA0321.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004470; AAG03710.1; -.
DR InterPro; IPR00286; His_deacetylse.
DR Pfam; PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 344 AA; 37883 MW; 8238194D8911D0CE CRC64;

Query Match 24.8%; Score 67.5; DB 16; Length 344;
Best Local Similarity 37.1%; Pred. No. 3.1;
Matches 13; Conservative 7; Mismatches 14; Indels 1; Gaps 1;

QY 1 QLNHEGFKLFATEATSDWL-NANNVPANVPVWPSQ 34
:::|:|:| | | | | | | | | | | | | | | | | | | |
Db 61 RVHSEGFVRFLQNAWQDWLATGRSHDMLPIANPTR 95

RESULT 10
Q8UG58
ID Q8UG58 PRELIMINARY; PRT; 863 AA.
AC Q8UG58;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE DNA-directed RNA polymerase.
GN ATU1183 OR AGR_C_2186.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kuttyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorllo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
```


DR InterPro; IPR000991; GAtase_1.
DR InterPro; IPR004362; MGS_like.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GAtase; 1.
DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR TIGRFAMS; TIGR01368; CPSaseIISmall; 1.
DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
SQ SEQUENCE 1505 AA; 164592 MW; 89C4A22285166CA2 CRC64;

Query Match 60.3%; Score 164; DB 13; Length 1505;
Best Local Similarity 64.7%; Pred. No. 1.1e-11;
Matches 33; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

Qy 1 QLNHEGKLFATEATSDWLNANNVPANVPVAPWPSQEQNP\$LSIRKLIRDG 51
|| |||||:||||| || ||:|||| |||||: | : || ||:|||| :||
Db 1380 QLKEGFKLYATEATSAWLCANDVPAIPVAPVTGMDGND\$LSIKRLISEG 1430

RESULT 5
Q92115 PRELIMINARY; PRT; 1502 AA.
AC Q92115; Q91438;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Carbamoyl phosphate synthetase III [glutamine] mitochondrial precursor
DE (EC 6.3.5.5) (Glutamine-dependent carbamoyl-phosphate synthetase III)
DE (GD-CPSase III).
GN CPSIII.
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 39-47.
RC TISSUE=LIVER;
RX MEDLINE=95018247; PubMed=7932737;
RA Hong J., Salo W.L., Lusty C.J., Anderson P.M.;
RT "Carbamyl phosphate synthetase III, an evolutionary intermediate in
RT the transition between glutamine-dependent and ammonia-dependent
RT carbamyl phosphate synthetases.";
RL J. Mol. Biol. 243:131-140(1994).
RN [2]
RP SEQUENCE OF 1-42 FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=97149029; PubMed=8995057;
RA Hong J., Salo W.L., Chen Y., Atkinson B.G., Anderson P.M.;
RT "The promoter region of the carbamoyl-phosphate synthetase III gene of
RT Squalus acanthias.";
RL J. Mol. Evol. 43:602-609(1996).
RN [3]
RP ENZYME ACTIVITY, AND REGULATION.
RC TISSUE=LIVER;
RX MEDLINE=80168997; PubMed=6245445;
RA Anderson P.M.;
RT "Glutamine- and N-acetylglutamate-dependent carbamoyl phosphate
RT synthetase in elasmobranchs.";
RL Science 208:291-293(1980).
RN [4]
RP ENZYME ACTIVITY, AND SUBUNITS.
RC TISSUE=LIVER;
RX MEDLINE=82053090; PubMed=7298655;
RA Anderson P.M.;
RT "Purification and properties of the glutamine- and N-acetyl-L-
RT glutamate-dependent carbamoyl phosphate synthetase from liver of
RT Squalus acanthias.";

J. Biol. Chem. 256:12228-12238(1981).
[5]
SUBCELLULAR LOCATION.
TISSUE=LIVER;
MEDLINE=82214067; PubMed=6123510;
Casey C.A., Anderson P.M.;
RT "Subcellular location of glutamine synthetase and urea cycle enzymes
RT in liver of spiny dogfish (Squalus acanthias).";
J. Biol. Chem. 257:8449-8453(1982).
CC -!- FUNCTION: INVOLVED IN THE UREA CYCLE OF UREOTELIC ANIMALS WHERE
CC THE ENZYME PLAYS AN IMPORTANT ROLE IN THE GLUTAMINE-DEPENDENT
CC FORMATION OF CARBAMOYL PHOSPHATE FOR UREA SYNTHESIS. UREA MAYBE
CC RETAINED IN TISSUES AND BODY FLUIDS AS A MECHANISM OF
CC OSMOREGULATION.
CC -!- CATALYTIC ACTIVITY: 2 ATP + GLUTAMINE + CO(2) + H(2)O = 2 ADP +
CC PHOSPHATE + GLUTAMATE + CARBAMOYL PHOSPHATE.
CC -!- ENZYME REGULATION: REQUIRES N-ACETYLGLUTAMATE AS AN ALLOSTERIC
CC ACTIVATOR.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LIVER. MAY NOT BE
CC EXPRESSED IN OTHER TISSUES.
CC -!- SIMILARITY: TO OTHER CARBAMOYL-PHOSPHATE SYNTHETASES. ALSO
CC CONTAINS A GLUTAMINE AMIDOTRANSFERASE-LIKE DOMAIN. THE CYSTEINE
CC RESIDUE WHICH IS ESSENTIAL FOR AMIDO TRANSFERASE ACTIVITY, IS
CC CONSERVED.
DR EMBL; L31362; AAA96435.1; -.
DR EMBL; U19771; AAB49032.1; -.
DR HSSP; P00968; ICS0.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; CPase_L.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005480; CPase_L_D3.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR002474; CPSase_sm_chain.
DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR000991; GAtase_1.
DR InterPro; IPR004362; MGS_like.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GAtase; 1.
DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR PRINTS; PR00096; GATASE.
DR TIGRFAMS; TIGR01368; CPSaseIISmall; 1.
DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
KW Ligase; Transit peptide; Mitochondrion; ATP-binding; Urea cycle.
FT TRANSIT 1 38 MITOCHONDRION.
FT CHAIN 39 1502 CARBAMOYL PHOSPHATE SYNTHETASE III
FT DOMAIN 39 219 [GLUTAMINE] MITOCHONDRIAL.
FT DOMAIN 220 411 ANTHRANILATE PHOSPHORIBOSYLTRANSFERASE-
FT DOMAIN 412 1502 LIKE.
FT NP_BIND 572 627 GLUTAMINE AMIDOTRANSFERASE-LIKE.
FT NP_BIND 719 769 CARBAMOYL-PHOSPHATE SYNTHETASE LARGE
FT NP_BIND 1114 1172 CHAIN.
FT NP_BIND 1259 1303 ATP (2) (POTENTIAL).
FT REPEAT 420 877 ATP (1) (POTENTIAL).
FT REPEAT 971 1411
FT ACT_SITE 294 294 GATASE (BY SIMILARITY).
SQ SEQUENCE 1502 AA; 165140 MW; CF444C222D99E2EE CRC64;

Query Match 59.9%; Score 163; DB 13; Length 1502;
Best Local Similarity 56.9%; Pred. No. 1.5e-11;
Matches 29; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

DT 01-FEB-1997 (TReMBLrel. 02, last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
DE Carbamoyl-phosphate synthetase III (EC 6.3.5.5).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97197794; PubMed=9045644;
RA Korte J.J., Salo W.L., Cabrera V.M., Wright P.A., Felskie A.K.,
RA Anderson P.M.;
RA "Expression of carbamoyl-phosphate synthetase III mRNA during the
RT early stages of development and in muscle of adult rainbow trout
RT (Oncorhynchus mykiss).";
RL J. Biol. Chem. 272:6270-6277(1997).
DR EMBL; U65893; AAC60207.1; -.
DR HSSP; P00968; ICS0.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; CPase_L.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005480; CPase_L_D3.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR002474; CPSase_sm_chain.
DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR000991; GATase_1.
DR InterPro; IPR004362; MGS_like.
DR Pfam; PF02789; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR PRINTS; PR00096; GATASE.
DR TIGRFAMS; TIGR01368; CPSaseIsmall; 1.
DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
KW Ligase.
SQ SEQUENCE 1518 AA; 166577 MW; F9D6331C710A0EBD CRC64;

Query Match 69.9%; Score 190; DB 13; Length 1518;
Best Local Similarity 72.5%; Pred. No. 5.9e-15;
Matches 37; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 QLHNEGKLFATEATSDWLNANNVPANVPVAVPWSQEGQNPSLSIRKLIRDG 51
||: ||||| ||||| ||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| : || || : || : |
Db 1378 QLNEEGKLFATEATSAWLSANDVPATPVAVPWSQEGGDASLPSIIRLINEG 1428

RESULT 3
O42433
ID O42433 PRELIMINARY; PRT; 1506 AA.
AC O42433;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
DE Carbamoyl-phosphate synthetase III (EC 6.3.5.5).
OS Micropterus salmoides (Largemouth bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Centrarchidae; Micropterus.
OX NCBI_TaxID=27706;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Kong H., Edberg D.D., Salo W.L., Korte J.J., Wright P.A.,
RA Anderson P.M.;

RT "Nitrogen Excretion and Expression of Carbamoyl-Phosphate Synthetase
RT III in Extra-Hepatic Tissues of Largemouth Bass (Micropterus
RT salmoides).";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF006491; AAB62566.1; -.
DR HSSP; P00968; ICS0.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; CPase_L.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005480; CPase_L_D3.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR002474; CPSase_sm_chain.
DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR000991; GATase_1.
DR InterPro; IPR004362; MGS_like.
DR Pfam; PF02789; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR PRINTS; PR00096; GATASE.
DR TIGRFAMS; TIGR01368; CPSaseIsmall; 1.
DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
DR PROSITE; PS00866; CPSASE_1; 1.
DR PROSITE; PS00867; CPSASE_2; 2.
KW Ligase.
SQ SEQUENCE 1506 AA; 165304 MW; 59439669AC8BB86D CRC64;

Query Match 62.9%; Score 171; DB 13; Length 1506;
Best Local Similarity 64.0%; Pred. No. 1.5e-12;
Matches 32; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 2 LHNEGKLFATEATSDWLNANNVPANVPVAVPWSQEGQNPSLSIRKLIRDG 51
| |||||:||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| : || || : || : |
Db 1382 LKEGFKLYATEATSAWLCANDVPATPVAVPTEKGGDTSLPSIKRLINEG 1431

RESULT 4
Q9W7F3
ID Q9W7F3 PRELIMINARY; PRT; 1505 AA.
AC Q9W7F3;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
DE Carbamoyl-phosphate synthetase III.
OS Alcolapia grahami.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Alcolapia.
OX NCBI_TaxID=87886;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99445540; PubMed=10514466;
RA Lindley T.E., Scheiderer C.L., Walsh P.J., Wood C.M., Bergman H.L.,
RA Bergman A.L., Laurent P., Wilson P., Anderson P.M.;
RT "Muscle as the primary site of urea cycle enzyme activity in an
RT alkaline lake-adapted tilapia, Oreochromis alcalicus grahami.";
RL J. Biol. Chem. 274:29858-29861(1999).
DR EMBL; AF119250; AAD43968.1; -.
DR HSSP; P00968; 1A9X.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; CPase_L.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005480; CPase_L_D3.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR002474; CPSase_sm_chain.
DR InterPro; IPR001317; CPS_GATase.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2003, 02:07:12 ; Search time 57.5 Seconds
(without alignments)
228.881 Million cell updates/sec

Title: US-09-585-077C-2_COPY_1380_1430
Perfect score: 272
Sequence: 1 QLNHEGFKLFATEATSDWLN.....PSQEGQNPSLSIRKLIRDG 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	266	97.8	Q8C196 mus musculus
2	190	69.9	P70087 oncorhynchus
3	171	62.9	O42433 micropterus
4	164	60.3	Q9W7F3 alcolapia g
5	163	59.9	Q92115 squalus aca
6	156.5	57.5	Q9PTM9 opsanus bet
7	80	29.4	Q18990 caenorhabdi
8	69	25.4	Q93937 emerice
9	67.5	24.8	Q9I6H0 pseudomonas
10	64	23.5	Q8UG58 agrobacteri
11	63.5	23.3	Q9W709 pleuronecte
12	63.5	23.3	Q98TI8 pleuronecte
13	63.5	23.3	Q9PW00 pleuronecte
14	63.5	23.3	Q14997 homo sapien
15	63.5	23.3	Q9ESY1 rattus norv
16	63	23.2	Q923L3 mus musculu

17	61.5	22.6	537	16	Q9A8M4	Q9a8m4 caulobacter
18	61	22.4	889	2	Q8KNY4	Q8kny4 xanthomonas
19	61	22.4	1065	16	Q8EZX0	Q8ezx0 leptospira
20	61	22.4	1113	16	Q8FT42	Q8ft42 corynebacte
21	61	22.4	2204	12	Q99FK6	Q99fk6 porcine tes
22	60.5	22.2	1360	12	O55253	O55253 murine hepa
23	60.5	22.2	1698	12	Q9YLR3	Q9ylr3 hepatitis e
24	60	22.1	336	16	Q97TA9	Q97ta9 streptococc
25	60	22.1	340	16	Q8DRN6	Q8drn6 streptococc
26	59	21.7	339	16	Q99ZV5	Q99zv5 streptococc
27	59	21.7	339	16	Q8K7M5	Q8k7m5 streptococc
28	59	21.7	354	16	Q8ESN9	Q8esn9 oceanobacil
29	59	21.7	375	10	Q8VXK9	Q8vxx9 fagus sylva
30	59	21.7	908	5	O18476	O18476 loligo peal
31	59	21.7	1020	4	Q9C0D4	Q9c0d4 homo sapien
32	59	21.7	1501	3	Q96VL9	Q96vl9 botrytis ci
33	58.5	21.5	700	16	Q9PFT1	Q9pft1 xylella fas
34	58.5	21.5	1068	16	Q8CXH7	Q8cxh7 oceanobacil
35	58.5	21.5	2921	5	Q9N973	Q9n973 leishmania
36	58	21.3	339	16	Q8P171	Q8p171 streptococc
37	58	21.3	540	3	Q00878	Q00878 nectria hae
38	57.5	21.1	97	16	Q8YVB1	Q8yyb1 anabaena sp
39	57.5	21.1	424	16	Q98HH3	Q98hh3 rhizobium l
40	57.5	21.1	508	4	Q8IZH4	Q8izh4 homo sapien
41	57.5	21.1	1181	4	Q9UM06	Q9um06 homo sapien
42	57.5	21.1	1220	4	Q9UM05	Q9um05 homo sapien
43	57.5	21.1	1220	4	O94979	O94979 homo sapien
44	57.5	21.1	1249	11	Q9Z2Q1	Q9z2q1 rattus norv
45	57	21.0	174	15	Q8QDL5	Q8qdl5 chimpanzee

ALIGNMENTS

RESULT 1			
Q8C196	Q8C196	PRELIMINARY;	PRT; 748 AA.
AC	Q8C196;		
DT	01-MAR-2003 (Tremblrel. 23, Created)		
DT	01-MAR-2003 (Tremblrel. 23, Last sequence update)		
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)		
DE	Weakly similar to carbamoyl-phosphate synthase (Fragment).		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Skin;		
RX	MEDLINE=22354683; PubMed=12466851;		
RA	The FANTOM Consortium,		
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;		
RT	"Analysis of the mouse transcriptome based on functional annotation of		
RT	60,770 full-length cDNAs."		
RL	Nature 420:563-573(2002).		
DR	EMBL; AK028683; BAC26064.1; -.		
FT	NON_TER		
SQ	SEQUENCE 748 AA; 83277 MW; AA818F3BAB82D9A1 CRC64;		
Query Match 97.8%; Score 266; DB 11; Length 748;			
Best Local Similarity 98.0%; Pred.No. 6.7e-25;			
Matches 50; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1 QLNHEGFKLFATEATSDWLNANNYPANVPVAVPSQEGQNPSLSIRKLIRDG 51		
Db	628 QLNHEGFKLFATEATSDWLNANNYPATPVAVPSQEGQNPSLSIRKLIRDG 678		
RESULT 2			
P70087	P70087	PRELIMINARY;	PRT; 1518 AA.
ID	P70087		
AC	P70087;		
DT	01-FEB-1997 (Tremblrel. 02, Created)		

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, last sequence update)

DT 28-FEB-2003 (Rel. 41, last annotation update)

DE URAL protein [includes: Glutamine-dependent carbamoyl-phosphate

DE synthase (EC 6.3.5.5); Aspartate carbamoyltransferase (EC 2.1.3.2)].

GN URAL OR SPAC22G7.06C.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RN [1]

RP SEQUENCE OF 22-2244 FROM N.A.

RC STRAIN=972;

RX MEDLINE=96020160; PubMed=8590465;

RA Lollier M., Jaquet L., Nedeva T., Lacroute F., Potier S.,

RA Souciet J.-L.;

RT "As in Saccharomyces cerevisiae, aspartate transcarbamoylase is

RT assembled on a multifunctional protein including a dihydroorotase-like

RT cryptic domain in Schizosaccharomyces pombe.";

RL Curr. Genet. 28:138-149(1995).

RN [2].

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgroos J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holroyd S., Hornsby T., Howarth S., Huckie E.J., Hunt S., Jagels K.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,

RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,

RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;

RT "The genome sequence of Schizosaccharomyces pombe";

RL Nature 415:871-880(2002).

CC -!- FUNCTION: THIS PROTEIN IS A "FUSION" PROTEIN ENCODING THREE

CC ENZYMATIC ACTIVITIES OF THE PYRIMIDINE PATHWAY (GATASE, CPSASE,

CC AND ATCASE) (BY SIMILARITY).

CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +

CC phosphate + L-glutamate + carbamoyl phosphate.

CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate = phosphate

CC + N-carbamoyl-L-aspartate.

CC -!- PATHWAY: Pyrimidine biosynthesis; first step.

CC -!- PATHWAY: Pyrimidine biosynthesis; second step.

CC -!- MISCELLANEOUS: GATASE (GLUTAMINE AMIDOTRANSFERASE) AND CPSASE

CC (CARBAMOYL PHOSPHATE SYNTHASE) FORM TOGETHER THE

CC GLUTAMINE-DEPENDENT CPSASE (GD-CPSASE) (EC 6.3.5.5).

CC -!- MISCELLANEOUS: IN EUKARYOTES EC 6.3.5.5 IS SYNTHESIZED BY TWO

CC PATHWAY-SPECIFIC (ARGININE AND PYRIMIDINE) UNDER SEPARATE CONTROL.

CC -!- SIMILARITY: THE CPSASE DOMAIN IS SIMILAR TO OTHER CPASES.

CC -!- SIMILARITY: IN THE CENTRAL SECTION; BELONGS TO THE DHOASE FAMILY.

CC DEFECTIVE DOMAIN.

CC -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X81841; CAA57433.1; -.

DR EMBL; Z54328; CAA91130.1; -.

DR PIR; S65074; S65074.

DR PIR; T11616; T11616.

DR HSSP; P00968; 1A9X.

DR GeneDB_SPombe; SPAC22G7.06c; -.

DR InterPro; IPR006220; Anth_synthII.

DR InterPro; IPR006130; Asp/Orn_Cotranf.

DR InterPro; IPR002082; Asp_carbmltransf.

DR InterPro; IPR006275; CarA_L_glu.

DR InterPro; IPR006274; CarA_small.

DR InterPro; IPR005483; CPase_L.

DR InterPro; IPR005479; CPase_L_D2.

DR InterPro; IPR005480; CPase_L_D3.

DR InterPro; IPR005481; CPase_L_N.

DR InterPro; IPR001317; CPS_GATase.

DR InterPro; IPR002474; CPSase_sm_chain.

DR InterPro; IPR000991; GATase_1.

DR InterPro; IPR004362; MGS_like.

DR InterPro; IPR006131; OTCace_O.

DR InterPro; IPR006132; OTCace_P.

DR Pfam; PF00289; CPSase_L_chain; 2.

DR Pfam; PF02786; CPSase_L_D2; 2.

DR Pfam; PF02787; CPSase_L_D3; 1.

DR Pfam; PF00988; CPSase_sm_chain; 1.

DR Pfam; PF00117; GATase; 1.

DR Pfam; PF02142; MGS; 1.

DR Pfam; PF00185; OTCace; 1.

DR Pfam; PF02729; OTCace_N; 1.

DR PRINTS; PR00097; ANTSNTHASEII.

DR PRINTS; PR00100; AOTCASE.

DR PRINTS; PR00098; CPSASE.

DR PRINTS; PR00099; CPSGATASE.

DR PRINTS; PR00096; GATASE.

DR TIGRFAMS; TIGR00670; asp_carb_tr; 1.

DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.

DR TIGRFAMS; TIGR01368; CPSaseII_small; 1.

DR TIGRFAMS; TIGR00097; CARBAMOYLTRANSFERASE; 1.

DR PROSITE; PS00442; GATASE_TYPE_I; 1.

DR PROSITE; PS00866; CPSASE_1; 2.

DR PROSITE; PS00867; CPSASE_2; 2.

KW Pyrimidine biosynthesis; Ligase; Transferase; Multifunctional enzyme.

FT DOMAIN 1 437 GATASE (GLUTAMINE AMIDOTRANSFERASE)

FT DOMAIN 438 477 (BY SIMILARITY).

FT DOMAIN 478 1514 LINKER (BY SIMILARITY).

FT DOMAIN 1515 1524 CPSASE (CARBAMOYL-PHOSPHATE SYNTHASE)

FT DOMAIN 1525 1853 (BY SIMILARITY).

FT DOMAIN 1854 1935 LINKER (BY SIMILARITY).

FT DOMAIN 1936 2244 DEFECTIVE DHOASE DOMAIN

FT ACT_SITE 338 338 (BY SIMILARITY).

FT ACT_SITE 422 422 GATASE (BY SIMILARITY).

FT ACT_SITE 424 424 GATASE (BY SIMILARITY).

FT CONFLICT 336 338 GIC -> RYF (IN REF. 1).

FT CONFLICT 1035 1039 CAVRA -> LQFAQ (IN REF. 1).

FT CONFLICT 1409 1410 EL -> DV (IN REF. 1).

FT CONFLICT 1975 1975 G -> E (IN REF. 1).

FT CONFLICT 2002 2002 G -> E (IN REF. 1).

SQ SEQUENCE 2244 AA; 248306 MW; 5700D153B50CD3E9 CRC64;

Query Match 23.5%; Score 64; DB 1; Length 2244;

Best Local Similarity 35.6%; Pred. No. 13;

Matches 16; Conservative 10; Mismatches 17; Indels 2; Gaps 2;

QY 1 QLHNEGFKLFATEATSDWLNANNVPANPVA-WPSQEQNP-SLSS 43

FT DOMAIN 367 1422 CPSASE (CARBAMOYL-PHOSPHATE SYNTHASE).
FT DOMAIN 367 909 CPSASE A.
FT DOMAIN 910 1422 CPSASE B.
FT DOMAIN 1423 1757 DHOASE (DIHYDROOROTASE).
FT DOMAIN 1758 1876 LINKER.
FT DOMAIN 1877 2185 ATCASE (ASPARTATE TRANSCARBAMYLASE).
FT ACT_SITE 236 236 GATASE (BY SIMILARITY).
FT ACT_SITE 322 322 GATASE (BY SIMILARITY).
FT ACT_SITE 324 324 GATASE (BY SIMILARITY).
FT METAL 1439 1439 ZINC (POTENTIAL).
FT METAL 1441 1441 ZINC (POTENTIAL).
FT CONFLICT 1185 1192 NNEIKVIE -> TMKSKLSN (IN REF. 1).
FT CONFLICT 1362 1362 L -> LF (IN REF. 1).
SQ SEQUENCE 2185 AA; 241741 MW; D0F32D879F06F7E5 CRC64;

Query Match 25.6%; Score 69.5; DB 1; Length 2185;
Best Local Similarity 35.3%; Pred. No. 2.7;
Matches 12; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

QY 6 GFKLFATEATSDWLNANNVPANVPANWPSQE-GQN 38
l: ||| : ||| : ||| : | : |||
D0 1316 GYTLFGTQGTADFYSENGVPVTQLNWDEEDLGEN 1349

RESULT 8
PYRL_YEAST STANDARD; PRT; 2214 AA.
AC P07259;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE URAl protein [Includes: Glutamine-dependent carbamoyl-phosphate
DE synthase (EC 6.3.5.5); Aspartate carbamoyltransferase (EC 2.1.3.2)].
GN UR2 OR YJL130C OR J0686.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 28583 / FL100;
RX MEDLINE=89378778; PubMed=2570735;
RA Souciet J.-L., Nagy M., le Gouar M., Lacroute F., Potier S.;
RT "Organization of the yeast UR2 gene: identification of a defective
RT dihydroorotase-like domain in the multifunctional carbamoylphosphate
RT synthetase-aspartate transcarbamylase complex.";
RL Gene 79:59-70(1989).
RN [2]
RP SEQUENCE OF 1-510 FROM N.A.
RX MEDLINE=87286375; PubMed=3039294;
RA Souciet J.-L., Potier S., Hubert J.-C., Lacroute F.;
RT "Nucleotide sequence of the pyrimidine specific carbamoyl phosphate
RT synthetase, a part of the yeast multifunctional protein encoded by
RT the UR2 gene.";
RL Mol. Gen. Genet. 207:314-319(1987).
RN [3]
RP SEQUENCE OF 1268-2214 FROM N.A.
RX MEDLINE=89255278; PubMed=2498313;
RA Nagy M., le Gouar M., Potier S., Souciet J.-L., Herve G.;
RT "The primary structure of the aspartate transcarbamylase region of
RT the UR2 gene product in Saccharomyces cerevisiae. Features involved
RT in activity and nuclear localization.";
RL J. Biol. Chem. 264:8366-8374(1989).
RN [4]
RP SEQUENCE OF 1-276 FROM N.A.
RC STRAIN=S288c / FY1679;
RX MEDLINE=97103775; PubMed=8948101;
RA Cziepluch C., Kordes E., Pujol A., Jauniaux J.-C.;
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X
RT reveals 19 open reading frames including UR2 (5' end), TRK1, PBS2,
RT SPT10, GCD14, RPE1, PHO86, NCA3, ASF1, CCT7, GZF3, two tRNA genes,
RT three remnant delta elements and a Ty4 transposon.";
RL Yeast 12:1471-1474(1996).

RN [5]
RP SEQUENCE OF 175-2214 FROM N.A.
RC STRAIN=S288c / FY1679;
RX MEDLINE=96408771; PubMed=8813765;
RA Katsoulou C., Tzermia M., Tavernarakis N., Alexandraki D.;
RT "Sequence analysis of a 40.7 kb segment from the left arm of yeast
RT chromosome X reveals 14 known genes and 13 new open reading frames
RT including homologues of genes clustered on the right arm of
RT chromosome XI.";
RL Yeast 12:787-797(1996).
CC -!- FUNCTION: THIS PROTEIN IS A "FUSION" PROTEIN ENCODING THREE
CC ENZYMATIC ACTIVITIES OF THE PYRIMIDINE PATHWAY (GATASE, CPSASE,
CC AND ATCASE).
CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + L-glutamate + carbamoyl phosphate.
CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate = phosphate
CC + N-carbamoyl-L-aspartate.
CC -!- PATHWAY: Pyrimidine biosynthesis; first step.
CC -!- PATHWAY: Pyrimidine biosynthesis; second step.
CC -!- MISCELLANEOUS: GATASE (GLUTAMINE AMIDOTRANSFERASE) AND CPSASE
CC (CARBAMOYL PHOSPHATE SYNTHASE) FORM TOGETHER THE
CC GLUTAMINE-DEPENDENT CPSASE (GD-CPSASE) (EC 6.3.5.5).
CC -!- MISCELLANEOUS: IN EUKARYOTES EC 6.3.5.5 IS SYNTHESIZED BY TWO
CC PATHWAY-SPECIFIC (ARGININE AND PYRIMIDINE) UNDER SEPARATE CONTROL.
CC -!- SIMILARITY: THE CPSASE DOMAIN IS SIMILAR TO OTHER CPSASES.
CC -!- SIMILARITY: IN THE CENTRAL SECTION; BELONGS TO THE DHOASE FAMILY.
CC DEFECTIVE DOMAIN.
CC -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M27174; AAA68280.1; ALT_INIT.
DR EMBL; X05553; CAA29068.1; -.
DR EMBL; J04711; AAA35198.1; -.
DR EMBL; Z49405; CAA89425.1; -.
DR EMBL; X87371; CAA60825.1; -.
DR EMBL; D28139; BAA05680.1; -.
DR PIR; S56911; QZBYU2.
DR HSSP; P00968; 1A9X.
DR SGD; S0003666; UR2.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0004070; F:aspartate carbamoyltransferase activity; IDA.
DR InterPro; IPR006220; Anth_synthII.
DR InterPro; IPR006130; Asp/Orn_COTranf.
DR InterPro; IPR002082; Asp_carbmltransf.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; Cpase_L.
DR InterPro; IPR005479; Cpase_L_D2.
DR InterPro; IPR005480; Cpase_L_D3.
DR InterPro; IPR005481; Cpase_L_N.
DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR002474; CPSase_sm_chain.
DR InterPro; IPR000991; GATase_1.
DR InterPro; IPR004362; MGS_like.
DR InterPro; IPR006131; OTCace_O.
DR InterPro; IPR006132; OTCace_P.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR Pfam; PF02142; MGS; 1.
DR Pfam; PF00185; OTCace; 1.
DR Pfam; PF02729; OTCace_N; 1.
DR PRINTS; PR00097; ANTSNTHASEII.
DR PRINTS; PR00100; AOTCASE.

DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPSASE.
DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
KW Pyrimidine biosynthesis; Ligase; Repeat; ATP-binding; Manganese;
KW Complete proteome.
FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.
FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 930 1071 ALLOSTERIC DOMAIN.
FT REPEAT 1 546
FT REPEAT 547 1071
FT NP_BIND 153 210 ATP (POTENTIAL).
FT NP_BIND 302 352 ATP (POTENTIAL).
FT METAL 284 284 MANGANESE 1 (BY SIMILARITY).
FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).
FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).
FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).
SQ SEQUENCE 1071 AA; 117649 MW; 195E33CFC5C6222C CRC64;

Query Match 25.6%; Score 69.5; DB 1; Length 1071;
Best Local Similarity 31.4%; Pred. No. 1.2;
Matches 16; Conservative 11; Mismatches 19; Indels 5; Gaps 1;

Qy 1 QLNHEGFKLFATERATSDWLNANNVPANVPANVPSQEQGNPSLSIRKLIRDG 51
Db 956 RFHAIQYNILATEGAGYLKEASIPAKVVGKIGDGNP-----LLDVIRNG 1001

RESULT 7
PYR1_DICDI STANDARD; PRT; 2185 AA.
AC P20054;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE protein PYR1-3 [includes: Glutamine-dependent carbamoyl-phosphate
DE synthase (EC 6.3.5.5); Aspartate carbamoyltransferase (EC 2.1.3.2);
DE Dihydroorotase (EC 3.5.2.3)].
GN PYR1-3.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE OF 1-467 AND 1175-2185 FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=89137111; PubMed=2917570;
RA Faure M., Camonis J.H., Jacquet M.;
RT "Molecular characterization of a Dictyostelium discoideum gene
RT encoding a multifunctional enzyme of the pyrimidine pathway.";
RL Eur. J. Biochem. 179:345-358(1989).
RN [2]
RP SEQUENCE OF 367-1408 FROM N.A.
RC STRAIN=AX2;
RX MEDLINE=92329976; PubMed=1627825;
RA Elgar G., Schofield J.P.;
RT "Carbamoyl phosphate synthetase (cpsase) in the PYR1-3 multigene of
RT Dictyostelium discoideum.";
RL DNA Seq. 2:219-226(1992).
CC -!- FUNCTION: THIS PROTEIN IS A "FUSION" PROTEIN ENCODING FOUR
CC ENZYMIC ACTIVITIES OF THE PYRIMIDINE PATHWAY (GATASE, CPSASE,
CC ATCASE AND DHOASE).
CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + L-glutamate + carbamoyl phosphate.
CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate = phosphate
CC + N-carbamoyl-L-aspartate.
CC -!- CATALYTIC ACTIVITY: (S)-dihydroorotate + H(2)O = N-carbamoyl-L-
CC aspartate.
CC -!- COFACTOR: DHOASE ACTIVITY REQUIRES A ZINC ATOM.
CC -!- ENZYME REGULATION: ALLOSTERICALLY REGULATED AND CONTROLLED BY
CC PHOSPHORYLATION. PRPP IS AN ACTIVATOR WHILE UMP IS AN INHIBITOR

CC OF THE CPSASE REACTION.
CC -!- PATHWAY: FIRST TO THIRD STEP IN PYRIMIDINE BIOSYNTHESIS.
CC -!- SUBUNIT: Homohexamer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- DEVELOPMENTAL STAGE: SEEN DURING GROWTH BUT NOT DURING
CC DEVELOPMENT.
CC -!- MISCELLANEOUS: GATASE (GLUTAMINE AMIDOTRANSFERASE) AND CPSASE
CC (CARBAMOYL PHOSPHATE SYNTHASE) FORM TOGETHER THE
CC GLUTAMINE-DEPENDENT CPSASE (GD-CPSASE) (EC 6.3.5.5).
CC -!- SIMILARITY: THE CPSASE DOMAIN IS SIMILAR TO OTHER CPASES.
CC -!- SIMILARITY: IN THE CENTRAL SECTION; BELONGS TO THE DHOASE FAMILY.
CC -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X14633; CAA32781.1; -.
DR EMBL; X14634; CAA32782.1; -.
DR EMBL; X55433; CAA39077.1; -.
DR PIR; S23738; S23738.
DR HSSP; P00479; 3CSU.
DR DictyDb; DD05019; pyr1-3.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR006130; Asp/Orn_Cotranf.
DR InterPro; IPR002082; Asp_carbmltransf.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; Cpase_L.
DR InterPro; IPR005479; Cpase_L_D2.
DR InterPro; IPR005480; Cpase_L_D3.
DR InterPro; IPR005481; Cpase_L_N.
DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR002474; CPSase_sm_chain.
DR InterPro; IPR002195; Dihydroorotase.
DR InterPro; IPR000991; GATase_1.
DR InterPro; IPR004362; MGS_like.
DR InterPro; IPR006131; OTCace_O.
DR InterPro; IPR006132; OTCace_P.
DR InterPro; IPR004722; PyrC_multi.
DR InterPro; IPR005847; Urease.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR Pfam; PF02142; MGS; 1.
DR Pfam; PF00185; OTCace; 1.
DR Pfam; PF02729; OTCace_N; 1.
DR PRINTS; PR00100; AOTCACE.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR PRINTS; PR00096; GATASE.
DR ProDom; PD000518; Urease; 1.
DR TIGRFAMS; TIGR00670; asp_carb_tr; 1.
DR TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.
DR TIGRFAMS; TIGR01368; CPSaseII_small; 1.
DR TIGRFAMS; TIGR00857; pyrC_multi; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
DR PROSITE; PS00442; GATASE_TYPE_I; 1.
DR PROSITE; PS00482; DIHYDROOROTASE_1; 1.
DR PROSITE; PS00483; DIHYDROOROTASE_2; 1.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
KW Pyrimidine biosynthesis; Ligase; Transferase; Hydrolase;
KW Allosteric enzyme; Zinc; Multifunctional enzyme.
FT DOMAIN 1 351 GATASE (GLUTAMINE AMIDOTRANSFERASE).
FT DOMAIN 352 366 LINKER.


```
CC similarity).
CC -!- SIMILARITY: BELONGS TO THE CARB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z81011; -; NOT_ANNOTATED_CDS.
CC EMBL; Z80108; -; NOT_ANNOTATED_CDS.
CC EMBL; AE007014; AAK45693.1; -.
CC PIR; A70990; A70990.
CC HSSP; P00968; 1CS0.
CC TIGR; MT1428; -.
CC TubercuList; Rvi384; -.
CC HAMAP; MF_01210; -; 1.
CC InterPro; IPR006275; CarA_L_glu.
CC InterPro; IPR005483; CPase_L.
CC InterPro; IPR005479; CPase_L_D2.
CC InterPro; IPR005480; CPase_L_D3.
CC InterPro; IPR005481; CPase_L_N.
CC InterPro; IPR004362; MGS_like.
CC Pfam; PF00289; CPSase_L_chain; 2.
CC Pfam; PF02786; CPSase_L_D2; 3.
CC Pfam; PF02787; CPSase_L_D3; 1.
CC Pfam; PF02142; MGS; 1.
CC PRINTS; PR00098; CPSASE.
CC TIGRFAMs; TIGR01369; CPSaseII_lrg; 1.
CC PROSITE; PS00866; CPSASE_1; 1.
CC PROSITE; PS00867; CPSASE_2; 2.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
KW ATP-binding; Manganese; Complete proteome.
FT DOMAIN 1 407 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 408 559 OLIGOMERIZATION DOMAIN.
FT DOMAIN 560 965 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 966 1115 ALLOSTERIC DOMAIN.
FT REPEAT 1 559
FT REPEAT 560 1115
FT NP_BIND 158 215 ATP (POTENTIAL).
FT NP_BIND 308 358 ATP (POTENTIAL).
FT METAL 290 290 MANGANESE 1 (BY SIMILARITY).
FT METAL 304 304 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 306 306 MANGANESE 2 (BY SIMILARITY).
FT METAL 843 843 MANGANESE 3 (BY SIMILARITY).
FT METAL 855 855 MANGANESE 3 (BY SIMILARITY).
SQ SEQUENCE 1115 AA; 118961 MW; 293F6A35FDE05981 CRC64;

Query Match 26.3%; Score 71.5; DB 1; Length 1115;
Best Local Similarity 35.2%; Pred. No. 0.69;
Matches 19; Conservative 14; Mismatches 16; Indels 5; Gaps 3;

QY 1 QLHNEGFKLFATEATSDWLNANNVPANPVA--WPSQEGQNPSLSIRKLIIRDG 51
Db 992 RLADLGRVRLATEGTAEMLRNRNGIPCDVYRKHFEPAPGR-PTMSAV-DAIRAG 1043

RESULT 6
CARB_BACSU
ID CARB_BACSU STANDARD; PRT; 1071 AA.
AC P25994;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carbamoyl-phosphate synthase, pyrimidine-specific, large chain
DE (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).
GN PYRAB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=91225016; PubMed=1709162;
RA Quinn C.L., Stephenson B.T., Switzer R.L.;
RT "Functional organization and nucleotide sequence of the Bacillus
RT subtilis pyrimidine biosynthetic operon.";
RL J. Biol. Chem. 266:9113-9127(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinios S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mael C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wiputt A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + L-glutamate + carbamoyl phosphate.
CC -!- COFACTOR: Binds 3 manganese ions per subunit (By similarity).
CC -!- PATHWAY: Pyrimidine biosynthesis; first step.
CC -!- SUBUNIT: Composed of two chains; the small (or glutamine) chain
CC promotes the hydrolysis of glutamine to ammonia, which is used by
CC the large (or ammonia) chain to synthesize carbamoyl phosphate.
CC -!- SIMILARITY: BELONGS TO THE CARB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M59757; AAA21270.1; -.
CC EMBL; Z99112; CAB13426.1; -.
CC PIR; F39845; F39845.
CC HSSP; P00968; 1JDB.
CC Subtilist; BG10716; pyrAB.
CC HAMAP; MF_01210; -; 1.
CC InterPro; IPR006275; CarA_L_glu.
CC InterPro; IPR005483; CPase_L.
CC InterPro; IPR005479; CPase_L_D2.
CC InterPro; IPR005480; CPase_L_D3.
CC InterPro; IPR005481; CPase_L_N.
CC InterPro; IPR004362; MGS_like.
CC Pfam; PF00289; CPSase_L_chain; 2.
CC Pfam; PF02786; CPSase_L_D2; 2.
CC Pfam; PF02787; CPSase_L_D3; 1.
```



```
FT REPEAT      416      873
FT REPEAT      967     1407
FT ACT_SITE    290      290      GATASE (BY SIMILARITY).
SQ SEQUENCE    1496 AA; 163767 MW; 2360BF05D21B059E CRC64;

Query Match
Best Local Similarity 66.0%; Score 179.5; DB 1; Length 1496;
Matches 35; Conservative 8; Mismatches 6; Indels 1; Gaps 1;

QY 2 LHNKGFKLFATEATSDWLNANNVPANVPVAVPWSQEGQNPSSLSSIRKLIRDG 51
   I :|||||:|||||:|||||: I :|||||:|||||: I :|||||:|||||:
Db 1378 LKDEGFKLYATEATADWLNANDITATPVAVPWSQEGQS-GPSSIIKLIKEG 1426

RESULT 4
CAB2_METJA STANDARD; PRT; 618 AA.
ID CAB2_METJA STANDARD; PRT; 618 AA.
AC Q58776;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carbamoyl-phosphate synthase large chain, C-terminal section
  (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).
GN CAB2 OR MJ1381.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT *Complete genome sequence of the methanogenic archaeon, Methanococcus
  jannaschii."
RL Science 273:1058-1073(1996).
CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
  phosphate + L-glutamate + carbamoyl phosphate.
CC -!- COFACTOR: Binds 3 manganese ions per subunit (By similarity).
CC -!- PATHWAY: Arginine biosynthesis.
CC -!- PATHWAY: Pyrimidine biosynthesis; first step.
CC -!- SUBUNIT: Composed of two chains; the small (or glutamine) chain
  promotes the hydrolysis of glutamine to ammonia, which is used by
  the large (or ammonia) chain to synthesize carbamoyl phosphate (By
  similarity).
CC -!- SIMILARITY: BELONGS TO THE CARB FAMILY. C-TERMINAL SECTION.
CC -!- CAUTION: Sequence of carb is split into two genes in M.jannaschii
  (MJ1378 and MJ1381).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67578; AAB99391.1; -.
CC FIR; D64472; D64472.
CC HSSP; P00968; 1BXR.
CC TIGR; MJ1381; -.
CC HAMAP; MF_01210; atypical; 1.
CC InterPro; IPR005483; CPase_L.
CC InterPro; IPR005479; CPase_L_D2.
CC InterPro; IPR005480; CPase_L_D3.
CC InterPro; IPR005481; CPase_L_N.
```

```
DR InterPro; IPR004362; MGS_like.
DR Pfam; PF00289; CPSase_L_chain; 1.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPSASE.
DR PROSITE; PS00866; CPSASE_1; 1.
DR PROSITE; PS00867; CPSASE_2; 1.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; ATP-binding;
  Manganese; Complete proteome.
FT DOMAIN 81 477 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 478 618 ALLOSTERIC DOMAIN.
FT METAL 358 358 MANGANESE 3 (BY SIMILARITY).
FT METAL 370 370 MANGANESE 3 (BY SIMILARITY).
SQ SEQUENCE 618 AA; 68453 MW; 8C3D4D9C192A954E CRC64;

Query Match 29.4%; Score 80; DB 1; Length 618;
Best Local Similarity 39.2%; Pred. No. 0.031;
Matches 20; Conservative 11; Mismatches 16; Indels 4; Gaps 2;

QY 1 QLNKGFKLFATEATSDWLNANNVPANVPVAVPWSQEGQNPSSLSSIRKLIRDG 51
   :|| || :||| | :| | :|| | :||| :||| :||| :||| :|||
Db 502 KLHELGFITYATEGTAKVLRENGIPAILV---KKISESPN-DNILKLRDGG 548

RESULT 5
CARB_MYCTU STANDARD; PRT; 1115 AA.
ID CARB_MYCTU STANDARD; PRT; 1115 AA.
AC P57689;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-
  phosphate synthetase ammonia chain).
GN CARB OR RV1384 OR MT1428 OR MTCY02Bi2.18 OR MTCY21B4.01.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
  complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
  laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
  phosphate + L-glutamate + carbamoyl phosphate.
CC -!- COFACTOR: Binds 3 manganese ions per subunit (By similarity).
CC -!- PATHWAY: Arginine biosynthesis.
CC -!- PATHWAY: Pyrimidine biosynthesis; first step.
CC -!- SUBUNIT: Composed of two chains; the small (or glutamine) chain
  promotes the hydrolysis of glutamine to ammonia, which is used by
  the large (or ammonia) chain to synthesize carbamoyl phosphate (By
```

CC	DR	EMBL; M12335; AAB59717.1; -.	
DR	EMBL; M11710; AAB59717.1; JOINED.		
DR	EMBL; M12318; AAB59717.1; JOINED.		
DR	EMBL; M12319; AAB59717.1; JOINED.		
DR	EMBL; M12320; AAB59717.1; JOINED.		
DR	EMBL; M12321; AAB59717.1; JOINED.		
DR	EMBL; M12322; AAB59717.1; JOINED.		
DR	EMBL; M12323; AAB59717.1; JOINED.		
DR	EMBL; M12324; AAB59717.1; JOINED.		
DR	EMBL; M12325; AAB59717.1; JOINED.		
DR	EMBL; M12326; AAB59717.1; JOINED.		
DR	EMBL; M12327; AAB59717.1; JOINED.		
DR	EMBL; M12328; AAB59717.1; JOINED.		
DR	EMBL; J02805; AAA40959.1; -.		
DR	PIR; A28481; SYRTCA.		
DR	HSSP; P00968; 1CS0.		
DR	InterPro; IPR006275; CarA_L_glu.		
DR	InterPro; IPR006274; CarA_small.		
DR	InterPro; IPR005483; Cpase_L.		
DR	InterPro; IPR005479; Cpase_L_D2.		
DR	InterPro; IPR005480; Cpase_L_D3.		
DR	InterPro; IPR005481; Cpase_L_N.		
DR	InterPro; IPR001317; CPS_GATase.		
DR	InterPro; IPR002474; CPSase_sm_chain.		
DR	InterPro; IPR000991; GATase_1.		
DR	InterPro; IPR004362; MGS_like.		
DR	Pfam; PF00289; CPSase_L_chain; 2.		
DR	Pfam; PF02786; CPSase_L_D2; 2.		
DR	Pfam; PF02787; CPSase_L_D3; 1.		
DR	Pfam; PF00988; CPSase_L_D3; 1.		
DR	Pfam; PF00117; CPSase_sm_chain; 1.		
DR	Pfam; PF02142; MGS; 1.		
DR	PRINTS; PR00098; CPSASE.		
DR	PRINTS; PR00099; CPSGATASE.		
DR	PRINTS; PR00096; GATASE.		
DR	TIGRFAMS; TIGR01369; CpsaseII_lrg; 1.		
DR	TIGRFAMS; TIGR01368; CpsaseIIsmall; 1.		
DR	PROSITE; PS00866; CPSASE_1; 2.		
DR	PROSITE; PS00867; CPSASE_2; 2.		
KW	Ligase; Allosteric enzyme; Repeat; Transit peptide; Mitochondrion;		
KW	ATP-binding; Urea cycle.		
FT	TRANSIT 1 38	MITOCHONDRION.	
FT	CHAIN 39 1500	CARBAMOYL-PHOSPHATE SYNTHASE [AMMONIA].	
FT	NP_BIND 571 626	ATP (2).	
FT	NP_BIND 718 768	ATP (1).	
FT	NP_BIND 1113 1171	ATP (2).	
FT	NP_BIND 1259 1302	ATP (1).	
FT	DOMAIN 39 219	ANTHRANILATE PHOSPHORIBOSYLTRANSFERASE	
FT		HOMOLOG.	
FT	DOMAIN 220 410	GLUTAMINE AMIDOTRANSFERASE-LIKE.	
FT	DOMAIN 411 1500	CARBAMOYL-PHOSPHATE SYNTHETASE LARGE	
FT		CHAIN.	
FT	REPEAT 419 876		
FT	REPEAT 970 1410		
SQ	SEQUENCE 1500 AA; 164579 MW; 038E8F893DE1C34D CRC64;		
Query Match 97.8%; Score 266; DB 1; Length 1500;			
Best Local Similarity 98.0%; Pred. No. 6.5e-25;			
Matches 50; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	QLHNEGKLFATEATSDWLNANNVPANVPVAPWPSQEGQNPSLSIRKLIRDG 51	
Db	1380	QLHNEGKLFATEATSDWLNANNVPATPVAPWPSQEGQNPSLSIRKLIRDG 1430	
RESULT 3			
CPSM_RANCA			
ID	CPSM_RANCA	STANDARD;	PRT; 1496 AA.
AC	Q91293;		
DT	15-JUL-1999	(Rel. 38, Created)	
DT	15-JUL-1999	(Rel. 38, Last sequence update)	
DT	28-FEB-2003	(Rel. 41, Last annotation update)	

DE	Carbamoyl-phosphate synthase [ammonia], mitochondrial precursor	
DE	(EC 6.3.4.16) (Carbamoyl-phosphate synthetase I) (CPSASE I).	
OS	Rana catesbeiana (Bull frog).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.	
OX	NCBI_TaxID=8400;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Liver;	
RX	MEDLINE=94216272; PubMed=8163471;	
RA	Helbing C.C., Atkinson B.G.;	
RT	"3,5,3'-Triiodothyronine-induced carbamyl-phosphate synthetase gene	
RT	expression is stabilized in the liver of Rana catesbeiana tadpoles	
RT	during heat shock.";	
RL	J. Biol. Chem. 269:11743-11750(1994).	
CC	-!- FUNCTION: INVOLVED IN THE UREA CYCLE OF UREOTELIC ANIMALS WHERE	
CC	THE ENZYME PLAYS AN IMPORTANT ROLE IN REMOVING EXCESS AMMONIA	
CC	FROM THE CELL (BY SIMILARITY).	
CC	-!- CATALYTIC ACTIVITY: 2 ATP + NH(3) + CO(2) + H(2)O = 2 ADP +	
CC	phosphate + carbamoyl phosphate.	
CC	-!- ENZYME REGULATION: REQUIRES N-ACETYLGLUTAMATE AS AN ALLOSTERIC	
CC	ACTIVATOR (BY SIMILARITY).	
CC	-!- SUBCELLULAR LOCATION: Mitochondrial.	
CC	-!- SIMILARITY: TO OTHER CARBAMOYL-PHOSPHATE SYNTHETASES. ALSO	
CC	CONTAINS A GLUTAMINE AMIDOTRANSFERASE-LIKE DOMAIN.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; U05193; AAA19016.1; -.	
DR	PIR; I51170; I51170.	
DR	HSSP; P00968; 1CS0.	
DR	InterPro; IPR006275; CarA_L_glu.	
DR	InterPro; IPR006274; CarA_small.	
DR	InterPro; IPR005483; Cpase_L.	
DR	InterPro; IPR005479; Cpase_L_D2.	
DR	InterPro; IPR005480; Cpase_L_D3.	
DR	InterPro; IPR005481; Cpase_L_N.	
DR	InterPro; IPR001317; CPS_GATase.	
DR	InterPro; IPR002474; CPSase_sm_chain.	
DR	InterPro; IPR000991; GATase_1.	
DR	InterPro; IPR004362; MGS_like.	
DR	Pfam; PF00289; CPSase_L_chain; 2.	
DR	Pfam; PF02786; CPSase_L_D2; 2.	
DR	Pfam; PF02787; CPSase_L_D3; 1.	
DR	Pfam; PF00988; CPSase_sm_chain; 1.	
DR	Pfam; PF00117; GATase; 1.	
DR	Pfam; PF02142; MGS; 1.	
DR	PRINTS; PR00098; CPSASE.	
DR	PRINTS; PR00099; CPSGATASE.	
DR	PRINTS; PR00096; GATASE.	
DR	TIGRFAMS; TIGR01369; CPSaseI_lrg; 1.	
DR	TIGRFAMS; TIGR01368; CPSaseI_small; 1.	
DR	PROSITE; PS00866; CPSASE_1; 2.	
DR	PROSITE; PS00867; CPSASE_2; 2.	
KW	Ligase; Allosteric enzyme; Repeat; Transit peptide; Mitochondrion;	
KW	ATP-binding; Urea cycle.	
FT	TRANSIT 1 33	MITOCHONDRION (BY SIMILARITY).
FT	CHAIN 34 1496	CARBAMOYL-PHOSPHATE SYNTHASE [AMMONIA].
FT	NP_BIND 568 623	ATP (2).
FT	NP_BIND 715 765	ATP (1).
FT	NP_BIND 1110 1168	ATP (2).
FT	NP_BIND 1256 1299	ATP (1).
FT	DOMAIN 34 215	ANTHRANILATE PHOSPHORIBOSYLTRANSFERASE
FT		HOMOLOG.
FT	DOMAIN 216 407	GLUTAMINE AMIDOTRANSFERASE-LIKE.
FT	DOMAIN 408 1496	CARBAMOYL-PHOSPHATE SYNTHETASE LARGE
FT		CHAIN.

CC symptoms are vomiting in infancy, protein intolerance,
CC intermittent ataxia, seizures, lethargy, and mental retardation.
CC -!- SIMILARITY: TO OTHER CARBAMOYL-PHOSPHATE SYNTHETASES. ALSO
CC CONTAINS A GLUTAMINE AMIDOTRANSFERASE-LIKE DOMAIN BUT THE CYSTEINE
CC RESIDUE WHICH IS ESSENTIAL FOR AMIDO TRANSFERASE ACTIVITY, IS
CC MISSING IN THIS PROTEIN SO THIS DOMAIN IS NO LONGER ACTIVE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90282; BAA14328.1; -.
DR EMBL; Y15793; CAA75785.1; -.
DR EMBL; AF154830; AAD38072.1; -.
DR PIR; JQ1348; JQ1348.
DR HSSP; P00968; 1A9X.
DR Genew; HGNC:2323; CPS1.
DR GK; P31327; -.
DR MIM; 237300; -.
DR GO; GO:0005739; C:mitochondrion; TAS.
DR GO; GO:0004087; F:carbamoyl-phosphate synthase (ammonia) acti. . .; TAS.
DR InterPro; IPR006275; CarA_L_glu.
DR InterPro; IPR006274; CarA_small.
DR InterPro; IPR005483; CPase_L.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005480; CPase_L_D3.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR001317; CPS_GATase.
DR InterPro; IPR002474; CPSase_sm_chain.
DR InterPro; IPR000991; GATase_1.
DR InterPro; IPR004362; MGS_like.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPSASE.
DR PRINTS; PR00099; CPSGATASE.
DR PRINTS; PR00096; GATASE.
DR TIGRFAMS; TIGR01369; CPSaseII_lrig; 1.
DR TIGRFAMS; TIGR01368; CPSaseII_small; 1.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
KW Ligase; Allosteric enzyme; Repeat; Transit peptide; Mitochondrion;
KW ATP-binding; Urea cycle; Polymorphism; Disease mutation.
FT TRANSIT 1 38 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 39 1500 CARBAMOYL-PHOSPHATE SYNTHASE [AMMONIA].
FT NP_BIND 571 626 ATP (2).
FT NP_BIND 718 768 ATP (1).
FT NP_BIND 1113 1171 ATP (2).
FT NP_BIND 1259 1302 ATP (1).
FT DOMAIN 39 219 ANTHRANILATE PHOSPHORIBOSYLTRANSFERASE
FT HOMOLOG.
FT DOMAIN 220 410 GLUTAMINE AMIDOTRANSFERASE-LIKE.
FT DOMAIN 411 1500 CARBAMOYL-PHOSPHATE SYNTHETASE LARGE
FT CHAIN.
FT REPEAT 419 876
FT REPEAT 970 1410
FT VARIANT 337 337
FT VARIANT 344 344
FT VARIANT 344 344
FT VARIANT 544 544
FT VARIANT 544 544
FT CONFLICT 111 111
FT CONFLICT 279 279
FT CONFLICT 338 338
FT CONFLICT 718 722
H -> R (in CPS1 deficiency).
/FTid=VAR_014077.
T -> A.
/FTid=VAR_006834.
T -> M (in CPS1 deficiency).
/FTid=VAR_006835.
A -> S (IN REF. 1).
R -> Q (IN REF. 1).
G -> C (IN REF. 1).
RLSRs--> KMSPN (IN REF. 1).

FT CONFLICT 729 729 A -> T (IN REF. 1).
FT CONFLICT 749 749 E -> G (IN REF. 1).
FT CONFLICT 1161 1162 EH -> AT (IN REF. 1).
FT CONFLICT 1204 1205 GD -> EN (IN REF. 1).
FT CONFLICT 1254 1254 I -> N (IN REF. 1).
FT CONFLICT 1266 1266 F -> S (IN REF. 1).
FT CONFLICT 1283 1283 M -> L (IN REF. 1).
FT CONFLICT 1303 1303 A -> V (IN REF. 1).
FT CONFLICT 1406 1406 T -> N (IN REF. 1).
SQ SEQUENCE 1500 AA; 164939 MW; E53A22D77563961D CRC64;
Query Match 97.8%; Score 266; DB 1; Length 1500;
Best local Similarity 98.0%; Pred. No. 6.5e-25;
Matches 50; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 QLNHEGFKLFATEATSDWLNANNVPANVPVAVPWSQEQNPSSLSSIRKLIRDG 51
|||||
Db 1380 QLNHEGFKLFATEATSDWLNANNVPATPVAVPWSQEQNPSSLSSIRKLIRDG 1430
RESULT 2
CPSM_RAT
ID CPSM_RAT STANDARD; PRT; 1500 AA.
AC P07756;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carbamoyl-phosphate synthase [ammonia], mitochondrial precursor
DE (EC 6.3.4.16) (Carbamoyl-phosphate synthetase I) (CPSASE I).
GN CPS1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85261323; PubMed=2991241;
RA Nyunoya H., Broglie K.E., Widgren E.E., Lusty C.J.;
RT "Characterization and derivation of the gene coding for mitochondrial
RT carbamyl phosphate synthetase I of rat.";
RL J. Biol. Chem. 260:9346-9356(1985).
RN [2]
RP SEQUENCE OF 1-42 FROM N.A.
RX MEDLINE=87280088; PubMed=3038878;
RA Lagace M., Howell B.W., Burak R., Lusty C.J., Shore G.C.;
RT "Rat carbamyl-phosphate synthetase I gene. Promoter sequence and
RT tissue-specific transcriptional regulation in vitro.";
RL J. Biol. Chem. 262:10415-10418(1987).
CC -!- FUNCTION: INVOLVED IN THE UREA CYCLE OF UREOTELIC ANIMALS WHERE
CC THE ENZYME PLAYS AN IMPORTANT ROLE IN REMOVING EXCESS AMMONIA
CC FROM THE CELL.
CC -!- CATALYTIC ACTIVITY: 2 ATP + NH(3) + CO(2) + H(2)O = 2 ADP +
CC phosphate + carbamoyl phosphate.
CC -!- ENZYME REGULATION: REQUIRES N-ACETYLGLUTAMATE AS AN ALLOSTERIC
CC ACTIVATOR.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- TISSUE SPECIFICITY: PRIMARILY IN THE LIVER AND SMALL INTESTINE.
CC -!- PTM: 50% OF THE MATURE PROTEIN THAT WAS ISOLATED HAD LEU 39 AS ITS
CC N-TERMINAL RESIDUE AND 50% HAD SER 40 SUGGESTING TWO ADJACENT
CC PROCESSING SITES. HOWEVER, THE POSSIBILITY OF PROTEOLYTIC REMOVAL
CC OF LEU 39 DURING THE ISOLATION OF THE ENZYME CANNOT BE EXCLUDED.
CC -!- SIMILARITY: TO OTHER CARBAMOYL-PHOSPHATE SYNTHETASES. ALSO
CC CONTAINS A GLUTAMINE AMIDOTRANSFERASE-LIKE DOMAIN BUT THE CYSTEINE
CC RESIDUE WHICH IS ESSENTIAL FOR AMIDO TRANSFERASE ACTIVITY, IS
CC MISSING IN THIS PROTEIN SO THIS DOMAIN IS NO LONGER ACTIVE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2003, 02:03:24 ; Search time 14 Seconds
(without alignments)
171.312 Million cell updates/sec

Title: US-09-585-077C-2_COPY_1380_1430
Perfect score: 272
Sequence: 1 OLHNEGFKLFATEATSDWLN.....PSQEGQNPSLSIRKLIRDG 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	266	97.8	1500	1	CPSM_HUMAN
2	266	97.8	1500	1	CPSM_RAT
3	179.5	66.0	1496	1	CPSM_RANCA
4	80	29.4	618	1	CAB2_METJA
5	71.5	26.3	1115	1	CARB_MYCTU
6	69.5	25.6	1071	1	CARB_BACSU
7	69.5	25.6	2185	1	PYR1_DICDI
8	65	23.9	2214	1	PYR1_YEAST
9	64	23.5	314	1	APAH_XANAC
10	64	23.5	452	1	PUR9_THEMA
11	64	23.5	1102	1	CARB_STRCO
12	64	23.5	2244	1	PYR1_SCHPO
13	60.5	22.2	1228	1	SLAP_BACST
14	59.5	21.9	199	1	PEXH_YEAST
15	58.5	21.5	1121	1	CARB_MYCLE
16	58	21.3	2242	1	PYR1_SQUAC
17	56.5	20.8	457	1	BAG4_MOUSE
18	56.5	20.8	811	1	RFX1_YEAST
19	56	20.6	521	1	YF92_MYCPN
20	56	20.6	591	1	ALU8_HUMAN
21	56	20.6	772	1	LMET_HUMAN
22	56	20.6	1113	1	CARB_CORGL
23	55.5	20.4	270	1	TCMP_STRGA
24	55	20.2	767	1	HYPE_SYNY3
25	55	20.2	786	1	BTKL_DROME
26	55	20.2	790	1	SMY2_YEAST
27	54.5	20.0	662	1	YFL8_YEAST
28	54.5	20.0	1034	1	BGAL_KLEPN
29	54.5	20.0	1092	1	CARB_ZYMMO
30	54.5	20.0	2225	1	PYR1_HUMAN
31	54.5	20.0	2225	1	PYR1_MESAU
32	54	19.9	441	1	PAC1_MOUSE
33	54	19.9	441	1	PAC1_RAT

34	54	19.9	444	1	PAC1_HUMAN
35	54	19.9	1076	1	CARB_ARCFU
36	54	19.9	1131	1	DNBI_HSV7J
37	54	19.9	1164	1	YAVL_XANCY
38	54	19.9	2280	1	COAC_SCHPO
39	53.5	19.7	323	1	VAL1_PASVK
40	53.5	19.7	423	1	Y37Q_RHIME
41	53.5	19.7	1080	1	CARB_XANAC
42	53	19.5	254	1	VSPA_SOYBN
43	53	19.5	291	1	S25K_SOYBN
44	53	19.5	423	1	EPD1_RALSO
45	53	19.5	488	1	PAC2_RAT

ALIGNMENTS

RESULT 1					
ID	CPSM_HUMAN	STANDARD;	PRT;	1500	AA.
AC	P31327; O43774;				
DT	01-JUL-1993 (Rel. 26, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Carbamoyl-phosphate synthase [ammonia], mitochondrial precursor				
DE	(EC 6.3.4.16) (Carbamoyl-phosphate synthetase I) (CPSase I).				
GN	CPS1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Liver;				
RX	MEDLINE=92084128; PubMed=1840546;				
RA	Haraguchi Y., Uchino T., Takiguchi M., Endo F., Mori M.,				
RA	Matsuda I.;				
RT	"Cloning and sequence of a cDNA encoding human carbamyl phosphate				
RT	synthetase I: molecular analysis of hyperammonemia.";				
RL	Gene 107:335-340(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A., VARIANT CPS1 DEFICIENCY MET-544, AND VARIANT				
RP	ALA-344.				
RC	TISSUE=Liver;				
RX	MEDLINE=98375696; PubMed=9711878;				
RA	Finckh U., Kohlschuetter A., Schaefer H., Sperhake K., Colombo J.-P.,				
RA	Gal A.;				
RT	"Prenatal diagnosis of carbamoyl phosphate synthetase I deficiency by				
RT	identification of a missense mutation in CPS1.";				
RL	Hum. Mutat. 12:206-211(1998).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RA	Summar M.;				
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	VARIANT CPS1 DEFICIENCY ARG-337.				
RX	MEDLINE=21367742; PubMed=11474210;				
RA	Aoshima T., Kajita M., Sekido Y., Kikuchi S., Yasuda I., Saheki T.,				
RA	Watanabe K., Shimokata K., Niwa T.;				
RT	"Novel mutations (H337R and 238-362del) in the CPS1 gene cause				
RT	carbamoyl phosphate synthetase I deficiency.";				
RL	Hum. Hered. 52:99-101(2001).				
CC	-!- FUNCTION: INVOLVED IN THE UREA CYCLE OF UREOTELIC ANIMALS WHERE				
CC	THE ENZYME PLAYS AN IMPORTANT ROLE IN REMOVING EXCESS AMMONIA				
CC	FROM THE CELL.				
CC	-!- CATALYTIC ACTIVITY: 2 ATP + NH(3) + CO(2) + H(2)O = 2 ADP +				
CC	phosphate + carbamoyl phosphate.				
CC	-!- ENZYME REGULATION: REQUIRES N-ACETYLGLUTAMATE AS AN ALLOSTERIC				
CC	ACTIVATOR.				
CC	-!- SUBCELLULAR LOCATION: Mitochondrial.				
CC	-!- TISSUE SPECIFICITY: PRIMARILY IN THE LIVER AND SMALL INTESTINE.				
CC	-!- DISEASE: Defects in CPS1 are the cause of an autosomal recessive				
CC	metabolic disorder that cause a type of hyperammonemia. Clinical				

Q9byl1	homo sapien
O28994	archaeoglob
P52339	human herpe
P14727	xanthomonas
P78820	schizosacch
Q00338	panicum str
Q52997	rhizobium m
P58942	xanthomonas
P15490	glycine max
P10742	glycine max
P58591	raistonia s
Q9qy17	rattus norv

Best Local Similarity 30.2%; Pred. No. 8.6;
Matches 19; Conservative 7; Mismatches 23; Indels 14; Gaps 2;
QY 3 HNEGKLFATEATSDWL---NANNVPANPVAMPSEQGNPSLSIR-----KLI 48
Db 221 HNTNIYVATKETLDWLATENSRLAPLSPVYLPVPPRPWTSPFRRGGYWSGRVRLRI 280
QY 49 RDG 51
Db 281 KTG 283

Search completed: September 13, 2003, 02:16:42
Job time : 28.5 secs

R;Denis-Duphil, M.; Lecaer, J.P.; Hardie, D.G.; Carrey, E.A.
Eur. J. Biochem. 193, 581-587, 1990
A;Title: Yeast carbamoyl-phosphate-synthetase - aspartate-transcarbamylase multidomain p
A;Reference number: S13358; MUID:91031508; PMID:1977585
A;Accession: S13358
A;Molecule type: protein
A;Residues: 1855-1882 <DEN>
R;Katsoulou, C.; Tzermia, M.; Tavernarakis, N.; Alexandraki, D.
Yeast 12, 787-797, 1996
A;Title: Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome X
of chromosome XI.
A;Reference number: S71643; MUID:96408771; PMID:8813765
A;Accession: S71667
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 175-2214 <KAW>
A;Cross-references: EMBL:X87371; NID:g854542; PIDN:CAA60825.1; PID:g854566
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
R;Nagy, M.; le Gouar, M.; Potier, S.; Souciet, J.L.; Herve, G.
J. Biol. Chem. 264, 8366-8374, 1989
A;Title: The primary structure of the aspartate transcarbamylase region of the URA2 gene
A;Reference number: A33820; MUID:89255278; PMID:2498313
A;Accession: S22790
A;Molecule type: DNA
A;Residues: 1268-1410,'S',1413-1581,'M',1583-1587,'K',1589-1591,'G',1593-1594,'A',1596-2
A;Cross-references: EMBL:J04711
C;Genetics:
A;Gene: SGD:URA2; MIPS:YJL130C
A;Cross-references: SGD:S0003666; MIPS:YJL130C
A;Map position: 10L
C;Function:
A;Description: ligase; methyltransferase
A;Pathway: pyrimidine nucleotide biosynthesis
C;Superfamily: rudimentary enzyme; aspartate/ornithine carbamoyltransferase homology; Ba
arbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homology; carbamoyl-phos
C;Keywords: ligase; methyltransferase; multifunctional enzyme; nucleus; phosphoprotein;
F;24-1487/Domain: carbamoyl-phosphate synthase (ammonia) homology <CPA>
F;24-404/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain homold
F;229-404/Domain: trpg homology <TRG>
F;440-1484/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homd
F;440-891/Domain: biotin carboxylase homology <BC1>
F;981-1424/Domain: biotin carboxylase homology <BC2>
F;1500-1825/Domain: Bacillus dihydroorotase homology <DHO>
F;1911-2210/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>
F;302/Active site: Cys #status predicted
Query Match 23.9%; Score 65; DB 1; Length 2214;
Best Local Similarity 45.8%; Pred. No. 20;
Matches 11; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY 1 QLNHEGFKLFATEATSDWLNANNV 24
Db 1382 KLYNMGYKLFATSGTADFLSEHGI 1405
RESULT 13
C72277
phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase - Thermotog
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: C72277
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: C72277
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-452 <ARN>
A;Cross-references: GB:AE001780; GB:AE000512; NID:g4981796; PIDN:AAD36324.1; PID:g498180
A;Experimental source: strain MSB8

C;Genetics:
A;Gene: Tml249
C;Superfamily: purH bifunctional enzyme
Query Match 23.5%; Score 64; DB 2; Length 452;
Best Local Similarity 33.3%; Pred. No. 4;
Matches 10; Conservative 12; Mismatches 8; Indels 0; Gaps 0;
QY 1 QLNHEGFKLFATEATSDWLNANNVPANPVA 30
Db 20 ELHEKGWEIWAASGTAKFLKSNGIEANDVS 49
RESULT 14
AE2722
DNA-directed RNA polymerase [imported] - Agrobacterium tumefaciens (strain C58, Dupon
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AE2722
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AE2722
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-863 <KUR>
A;Cross-references: GB:AE008688; PIDN:AAL42195.1; PID:gl7739586; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atul183
A;Map position: circular chromosome
C;Superfamily: phage T7 DNA-directed RNA polymerase
Query Match 23.5%; Score 64; DB 2; Length 863;
Best Local Similarity 30.2%; Pred. No. 8.6;
Matches 19; Conservative 7; Mismatches 23; Indels 14; Gaps 2;
QY 3 HNEGFKLFATEATSDWL--NANNVPANVPWPSQEQNPSSLSSIR-----KLI 48
Db 221 HNTNIVVATKETLDWLATENSRLAPLSPVYLFTLVPPRPWTSPFRGGYWSGRVRLRI 280

QY 49 RDG 51
Db 281 KTG 283

RESULT 15
B97504
RNA polymerase [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: B97504
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: B97504
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-863 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK86987.1; PID:gl5156227; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C.2186
A;Map position: circular chromosome
C;Superfamily: phage T7 DNA-directed RNA polymerase
Query Match 23.5%; Score 64; DB 2; Length 863;

A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: A69686
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1071 <KUN>
A;Cross-references: GB:Z99112; GB:AL009126; NID:g2633902; PIDN:CAB13426.1; PID:g2633925
A;Experimental source: strain 168
C;Genetics:
A;Gene: pyrAB
C;Function:
A;Pathway: glutamate metabolism; pyrimidine nucleotide biosynthesis
C;Superfamily: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain; biotin
C;Keywords: ligase; pyrimidine nucleotide biosynthesis
F;9-1047/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homolog
F;9-467/Domain: biotin carboxylase homology <BC1>
F;553-985/Domain: biotin carboxylase homology <BC2>

Query Match 25.6%; Score 69.5; DB 2; Length 1071;
Best Local Similarity 31.4%; Pred. No. 2.3;
Matches 16; Conservative 11; Mismatches 19; Indels 5; Gaps 1;

QY 1 QLHNEGFKLFATEATSDWLNANNVPANPVAVPWSQEQNPSSLIRKLIRDG 51
Db 956 RFHAIGNILATEGTAGYLKEASIPAKVVGKIGQDGN-----LLDVIRNG 1001

RESULT 10
QZDOP3
pyrimidine synthesis protein PYR1-3 - slime mold (Dictyostelium discoideum) (fragments)
N;Contains: aspartate carbamoyltransferase (EC 2.1.3.2); carbamoyl-phosphate synthase (9
C;Species: Dictyostelium discoideum
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 07-Aug-1998
C;Accession: S02800
R;Faure, M.; Camonis, J.H.; Jacquet, M.
Eur. J. Biochem. 179, 345-358, 1989
A;Title: Molecular characterization of a Dictyostelium discoideum gene encoding a multifu
A;Reference number: S02800; MUID:89137111; PMID:2917570
A;Accession: S02800
A;Molecule type: DNA
A;Residues: 1-467;468-1481 <FAU>
A;Cross-references: EMBL:X14633
A;Experimental source: strain AX3
C;Genetics:
A;Gene: PYR1-3
C;Superfamily: rudimentary enzyme; aspartate/ornithine carbamoyltransferase homology; Ba
arbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homology; carbamoyl-phos
C;Keywords: hydrolase; ligase; methyltransferase; multifunctional enzyme; pyrimidine nuc
F;1-707/Domain: carbamoyl-phosphate synthase (ammonia) homology (fragments) <CPA>
F;1-340/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain homolog
F;158-340/Domain: trpG homology <TRG>
F;357-704/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homol
F;367-467/Domain: biotin carboxylase homology (fragment) <BC1>
F;468-651/Domain: biotin carboxylase homology (fragment) <BC2>
F;721-1066/Domain: Bacillus dihydroorotase homology <DHO>
F;1179-1477/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>
F;236/Active site: Cys #status predicted

Query Match 25.6%; Score 69.5; DB 1; Length 1481;
Best Local Similarity 35.3%; Pred. No. 3.3;
Matches 12; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

QY 6 GFKLFATEATSDWLNANNVPANPVAVPWSQE-GQN 38
Db 611 GYTLFGTGTADFYSENGVPTQLNWDDEEDIGEN 644

RESULT 11
B83605
probable acetylpolymaine aminohydrolase PA0321 [imported] - Pseudomonas aeruginosa (stra
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: B83605
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of pseudomonas aeruginosa PA01, an opportunistic pa
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: B83605
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-344 <STO>
A;Cross-references: GB:AE004470; GB:AE004091; NID:g9946164; PIDN:AAG03710.1; GSPDB:GN
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA0321

Query Match 24.8%; Score 67.5; DB 2; Length 344;
Best Local Similarity 37.1%; Pred. No. 1;
Matches 13; Conservative 7; Mismatches 14; Indels 1; Gaps 1;

QY 1 QLHNEGFKLFATEATSDWL-NANNVPANPVAVPSQ 34
Db 61 RVHSEGFVRFLQNAWQDWLATGRSHDMLPIAWPTR 95

RESULT 12
QZBYU2
pyrimidine synthesis protein URA2 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein J0686; protein YJL130c
N;Contains: aspartate carbamoyltransferase (EC 2.1.3.2); carbamoyl-phosphate synthase
C;Species: Saccharomyces cerevisiae
C;Date: 30-Sep-1991 #sequence_revision 08-Sep-1995 #text_change 12-Nov-1999
C;Accession: S56911; S56912; S05767; S05859; S55182; S13358; S71667; S22790; A33820;
R;Cziepluch, C.; Kordes, E.; Pujol, A.; Jauniaux, J.C.
submitted to the Protein Sequence Database, September 1995
A;Reference number: S56891
A;Accession: S56911
A;Molecule type: DNA
A;Residues: 1-276 <CZI>
A;Cross-references: EMBL:Z49405; GSPDB:GN00010; MIPS:YJL130c
R;Katsoulou, C.; Tzermia, M.; Alexandraki, D.
submitted to the Protein Sequence Database, September 1995
A;Reference number: S56912
A;Accession: S56912
A;Molecule type: DNA
A;Residues: 175-2214 <KAT>
A;Cross-references: EMBL:Z49405; GSPDB:GN00010; MIPS:YJL130c
R;Souciet, J.L.; Nagy, M.; le Gouar, M.; Lacroute, F.; Potier, S.
Gene 79, 59-70, 1989
A;Title: Organization of the yeast URA2 gene: identification of a defective dihydroor
A;Reference number: S05766; MUID:89378778; PMID:2570735
A;Accession: S05767
A;Molecule type: DNA
A;Residues: 1-122, 'A', 124-249, 'RI', 250, 'SCSMD', 258-269, 'Y', 271-312, 'VQ', 315-371, 'RF',
A', 1596-2214 <SOU1>
A;Cross-references: EMBL:M27174
R;Souciet, J.L.; Potier, S.; Hubert, J.C.; Lacroute, F.
Mol. Gen. Genet. 207, 314-319, 1987
A;Title: Nucleotide sequence of the pyrimidine specific carbamoyl phosphate synthetas
A;Reference number: S05859; MUID:87286375; PMID:3039294
A;Accession: S05859
A;Molecule type: DNA
A;Residues: 1-85, 'D', 87-122, 'A', 124-249, 'RI', 250, 'SCSMD', 258-269, 'Y', 271-312, 'VQ', 315
506 <SOU2>
A;Cross-references: EMBL:X05553; NID:g4760; PIDN:CAA29068.1; PID:g4761
A;Note: the authors translated the codon GAC for residue 85 as His
R;Katsoulou, C.; Tzermia, M.; Alexandraki, D.
submitted to the EMBL Data Library, May 1995
A;Description: The complete sequence of a 40.7 kb segment located on the left arm of
yeast hypothetical proteins.
A;Reference number: S55159
A;Accession: S55182
A;Molecule type: DNA
A;Residues: 175-2214 <KA2>
A;Cross-references: EMBL:X87371; NID:g854542; PIDN:CAA60825.1; PID:g854566

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2003, 02:11:17 ; Search time 23.5 Seconds
(without alignments)
208.707 Million cell updates/sec

Title: US-09-585-077C-2_COPY_1380_1430
Perfect score: 272
Sequence: 1 QLHNEGFKLFATEATSDWLN.....PSQEQNPSSLIRKLIRDG 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	272	100.0	1500	1 JQ1348	carbamoyl-phosphat
2	266	97.8	1500	1 SYRTCA	carbamoyl-phosphat
3	179.5	66.0	1496	2 I51170	carbamyl phosphate
4	163	59.9	1502	2 S53602	carbamoyl-phosphat
5	80	29.4	618	2 D64472	carbamoyl-phosphat
6	80	29.4	2198	2 T20371	hypothetical prote
7	71.5	26.3	1115	2 A70990	carbamoyl-phosphat
8	69.5	25.6	1042	2 S23738	pyr1-3 protein - s
9	69.5	25.6	1071	2 F39845	carbamoyl-phosphat
10	69.5	25.6	1481	1 Q2DOP3	pyrimidine synthes
11	67.5	24.8	344	2 B83605	probable acetylpol
12	65	23.9	2214	1 Q2BYU2	pyrimidine synthes
13	64	23.5	452	2 C72277	phosphoribosylamin
14	64	23.5	863	2 AE2722	DNA-directed RNA p
15	64	23.5	863	2 B97504	RNA polymerase [im
16	64	23.5	2223	2 S65074	pyrimidine synthes
17	64	23.5	2244	2 T11616	carbamoyl-phosphat
18	61.5	22.6	537	2 B87414	RNA polymerase sig
19	60.5	22.2	1228	2 I40468	surface layer prot
20	60	22.1	336	2 G95003	membrane protein [
21	60	22.1	340	2 B97876	conserved hypothet
22	59.5	21.9	199	2 S50717	probable membrane
23	58.5	21.5	700	2 B82788	metallopeptidase X
24	58.5	21.5	1129	2 H86975	probable carbamoyl
25	58	21.3	2242	2 A57541	pyrimidine synthes
26	57.5	21.1	97	2 AI1923	hypothetical prote
27	57.5	21.1	1249	2 T14150	vesicle associated
28	56.5	20.8	277	2 AE1182	sugar ABC transpor
29	56.5	20.8	277	2 AF1539	sugar ABC transpor

30	56.5	20.8	715	2 S54628	hypothetical prote
31	56.5	20.8	771	2 S51421	hypothetical prote
32	56	20.6	190	2 G84018	hypothetical prote
33	56	20.6	347	2 S55629	hypothetical prote
34	56	20.6	415	2 AE2237	hypothetical prote
35	56	20.6	521	2 S62794	probable lipoprote
36	55.5	20.4	271	2 C47127	tetracenomycin C s
37	55	20.2	297	2 T18637	hypothetical prote
38	55	20.2	492	2 F86263	hypothetical prote
39	55	20.2	767	2 S76302	hypothetical prote
40	55	20.2	790	2 S27458	SMY2 protein - yea
41	54.5	20.0	155	2 S77384	hypothetical prote
42	54.5	20.0	334	2 G75069	cytochrome-c3 hydr
43	54.5	20.0	546	2 T24679	hypothetical prote
44	54.5	20.0	662	2 S56303	probable membrane
45	54.5	20.0	673	4 F40201	artifact-warning s

ALIGNMENTS

RESULT 1

JQ1348
carbamoyl-phosphate synthase (ammonia) (EC 6.3.4.16) precursor - human
N;Alternate names: carbamyl phosphate synthetase I; carbon-dioxide-ammonia ligase
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1992 #sequence_revision 02-Jun-1994 #text_change 16-Jun-2000
C;Accession: JQ1348
R;Haraguchi, Y.; Uchino, T.; Takiguchi, M.; Endo, F.; Mori, M.; Matsuda, I.
Gene 107, 335-340, 1991
A;Title: Cloning and sequence of a cDNA encoding human carbamyl phosphate synthetase
A;Reference number: JQ1348; MUID:92084128; PMID:1840546
A;Accession: JQ1348
A;Molecule type: mRNA
A;Residues: 1-1500 <HAR>
A;Cross-references: DDBJ:D90282; NID:g219552; PIDN:BAAL4328.1; PID:g219553
C;Comment: This is the first enzyme of the urea cycle; it catalyzes the synthesis of
C;Genetics:
A;Gene: GDB:CPS1
A;Cross-references: GDB:119799; OMIM:237300
A;Map position: 2q33-2q36
C;Superfamily: carbamoyl-phosphate synthase (ammonia); biotin carboxylase homology; c
omology; carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain homology; t
C;Keywords: ATP; ligase; urea cycle
F;1-38/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;39-1500/Product: carbamoyl-phosphate synthase (ammonia) #status predicted <CAR>
F;48-1479/Domain: carbamoyl-phosphate synthase (ammonia) homology <CPA>
F;48-395/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain hom
F;220-395/Domain: trpg homology <TRG>
F;424-1476/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain h
F;424-886/Domain: biotin carboxylase homology <BC1>
F;975-1426/Domain: biotin carboxylase homology <BC2>

Query Match 100.0%; Score 272; DB 1; Length 1500;
Best Local Similarity 100.0%; Pred. No. 1.2e-25;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLHNEGFKLFATEATSDWLNANNVPANPVANVPVAVWPSQEQNPSSLIRKLIRDG 51
|||||
Db 1380 QLHNEGFKLFATEATSDWLNANNVPANPVANVPVAVWPSQEQNPSSLIRKLIRDG 1430

RESULT 2

SYRTCA
carbamoyl-phosphate synthase (ammonia) (EC 6.3.4.16) I precursor - rat
N;Alternate names: carbamyl phosphate synthetase I; carbon-dioxide-ammonia ligase
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 31-Dec-2000
C;Accession: A28481; A23580
R;Lagace, M.; Howell, B.W.; Burak, R.; Lusty, C.J.; Shore, G.C.
J. Biol. Chem. 262, 10415-10418, 1987
A;Title: Rat carbamyl-phosphate synthetase I gene. Promoter sequence and tissue-speci
A;Reference number: A28481; MUID:87280088; PMID:3038878

PT Misc-difference 847
FT /label= Unknown
FT /note= "Xaa can be any amino acid"
XX
PN WO200210199-A2.
XX
PD 07-FEB-2002.
XX
PF 24-JUL-2001; 2001WO-US23232.
XX
PR 02-AUG-2000; 2000US-222504P.
PR 28-NOV-2000; 2000US-0728787.
XX
PA (AMGE-) AMGEN INC.
XX
XX Welcher AA, Elliott GS;
PI
XX
DR WPI; 2002-303934/34.
DR N-PSDB; AAD33319.
XX
PT Nucleic acid encoding a novel C3b/C4b Complement Receptor-like nucleic
PT acid molecule, useful for treating, preventing and diagnosing
PT rheumatoid arthritis, psoriatic arthritis, inflammatory arthritis, and
PT multiple sclerosis -
XX
PS Claim 13; Fig 3; 251pp; English.
XX
CC The invention relates to a nucleic acid encoding a novel C3b/C4b
CC complement receptor (CR)-like nucleic acid molecule. The C3b/C4b CR-like
CC polypeptide and nucleic acid molecules may be used to treat, prevent,
CC ameliorate, diagnose and/or detect diseases such as immune system
CC disorders such as rheumatoid arthritis, psoriatic arthritis,
CC inflammatory arthritis, osteoarthritis, inflammatory joint disease,
CC autoimmune disease, multiple sclerosis, lupus, inflammatory bowel
CC disease, transplant rejection, nervous system disorders (e.g. Alzheimer's
CC disease), ischaemic conditions, metabolic disorders (e.g. obesity and
CC diabetes) and infertility. The invention is useful in gene therapy. The
CC present sequence is rat C3b/C4b complement receptor like protein.
XX
SQ Sequence 3095 AA;

Query Match 22.4%; Score 61; DB 23; Length 3095;
Best Local Similarity 40.5%; Pred. No. 1.9e+02;
Matches 17; Conservative 3; Mismatches 14; Indels 8; Gaps 2;

QY 4 NEGFKLFATEATSD-----WLNANNVP-ANPVAWPSQEGQ 37
| | | | | | : : | | | | | | | | | | | | | | | |
Db 2053 NEGFKLDASQQATAVCQEDGLWSNRGKPPPTCKPVPSPSIEGQ 2094

Search completed: September 13, 2003, 02:12:44
Job time : 55 secs

Best Local Similarity 37.0%; Pred. No. 0.82; Matches 17; Conservative 1; Mismatches 12; Indels 16; Gaps 2;

QY 18 WLNANNVPANPVAW-----PSQEGQNPSLSSIRKLIRDG 51
 || | | | | | | | | | |
Db 13 WL-----TPVIPVLWDPKADSGIRDOPSQHGETPSLKLIEKLACHG 54

RESULT 9
ABG09932
ID ABG09932 standard; Protein; 692 AA.
XX
AC ABG09932;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #9923.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS74119.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 40291; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 692 AA;
SQ
Query Match 23.3%; Score 63.5; DB 22; Length 692;
Best Local Similarity 37.8%; Pred. No. 12;
Matches 14; Conservative 7; Mismatches 9; Indels 7; Gaps 2;

QY 1 QLNHEGFKLFATEATSDWLNANNVPANPVAWPSEQG 37
 :|:|:|:|:|:| | |:|:|:| | | |
Db 232 ELHHKGFKLWFDELIGLWVSQNLPL----QW---EGQ 261

RESULT 10
ABG09933
ID ABG09933 standard; Protein; 1798 AA.
XX
AC ABG09933;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #9924.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS74120.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 40292; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1798 AA;
SQ
Query Match 23.3%; Score 63.5; DB 22; Length 1798;
Best Local Similarity 37.8%; Pred. No. 43;
Matches 14; Conservative 7; Mismatches 9; Indels 7; Gaps 2;
QY 1 QLNHEGFKLFATEATSDWLNANNVPANPVAWPSEQG 37

CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 344 AA;

Query Match 24.8%; Score 67.5; DB 22; Length 344;
Best Local Similarity 37.1%; Pred. No. 1.5;
Matches 13; Conservative 7; Mismatches 14; Indels 1; Gaps 1;

QY 1 QLHNEGFKLFATEATSDWL-NANNVPANVPVAPSQ 34
Db 61 RVHSEGFVRFQLQAWQDWLATGRSHDMLPIAWPTR 95

RESULT 7
AAO13747
ID AAO13747 standard; Protein; 43 AA.
XX
AC AAO13747;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 27639.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.

XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.

XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI93678.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -

XX
PS Claim 20; SEQ ID NO 27639; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 43 AA;

Query Match 24.3%; Score 66; DB 22; Length 43;
Best Local Similarity 34.9%; Pred. NO. 0.16;
Matches 15; Conservative 6; Mismatches 18; Indels 4; Gaps 1;

QY 1 QLHNEGFKLFATEATSDWLNANNVPANVPVAPSQEQGNPSLSS 43
Db 5 QIHKLHFKKFRILGWTWWL-----TPVIPAPWEAEGRSPSPSS 43

RESULT 8
AAO09946
ID AAO09946 standard; Protein; 94 AA.
XX
AC AAO09946;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 23838.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.

XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.

XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI89877.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -

XX
PS Claim 20; SEQ ID NO 23838; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

XX
SQ Note: The sequence data for this patent did not form part of the printed
SQ specification, but was obtained in electronic format directly from WIPO
SQ at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 94 AA;

Query Match 23.5%; Score 64; DB 22; Length 94;

OS Homo sapiens.
XX WO200073322-A1.
PN
XX
XX 07-DEC-2000.
PD
XX
PF 01-JUN-2000; 2000WO-US15079.
XX
XX 01-JUN-1999; 99US-0323472.
PR
XX (UYVA-) UNIV VANDERBILT.
PA
XX Summar ML, Christman BW;
PI
XX
XX WPI; 2001-049926/06.
DR
XX
XX
PT Detecting susceptibility to suboptimal urea cycle function, e.g. bone
PT marrow transplant toxicity, comprises identifying a polymorphism in the
PT gene for carbamylphosphate synthase -
XX
XX Claim 57; Page 162-165; 171pp; English.
PS
XX The present invention relates to screening for susceptibility to
CC suboptimal urea cycle function or to bone marrow transplant toxicity
CC by detecting a polymorphism in the carbamyl phosphate synthase I
CC (CPSI) gene. The method is used to detect subjects at risk of
CC hepatitis, sclerosis, pulmonary hypertension and bone marrow
CC transplant toxicity. These conditions can be treated or prevented
CC by administration of a nitric oxide precursor or by gene
CC therapy (administration of sequences that encode CPSI).
XX
XX Sequence 1500 AA;
SQ

Query Match 100.0%; Score 272; DB 22; Length 1500;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLNHEGFKLFATEATSDWLNANNVPANPVAVPWSQEGQNPSLSIRKLIRDG 51
|||||
Db 1380 QLNHEGFKLFATEATSDWLNANNVPANPVAVPWSQEGQNPSLSIRKLIRDG 1430

RESULT 5
AAB49223
ID AAB49223 standard; protein; 1500 AA.
XX
AC AAB49223;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human T1405 carbamyl phosphate synthase I DNA.
XX
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy.
XX
XX Homo sapiens.
OS
XX
XX WO200073322-A1.
PN
XX
XX 07-DEC-2000.
PD
XX
XX 01-JUN-2000; 2000WO-US15079.
PF
XX
XX 01-JUN-1999; 99US-0323472.
PR
XX (UYVA-) UNIV VANDERBILT.
PA
XX Summar ML, Christman BW;
PI
XX
XX WPI; 2001-049926/06.
DR
XX
XX Detecting susceptibility to suboptimal urea cycle function, e.g. bone
PT marrow transplant toxicity, comprises identifying a polymorphism in the

PT gene for carbamylphosphate synthase -
XX
XX Disclosure; Fig 11; 171pp; English.
PS
XX
XX The present invention relates to screening for susceptibility to
CC suboptimal urea cycle function or to bone marrow transplant toxicity
CC by detecting a polymorphism in the carbamyl phosphate synthase I
CC (CPSI) gene. The method is used to detect subjects at risk of
CC hepatitis, sclerosis, pulmonary hypertension and bone marrow
CC transplant toxicity. These conditions can be treated or prevented
CC by administration of a nitric oxide precursor or by gene
CC therapy (administration of sequences that encode CPSI).
XX
XX Sequence 1500 AA;
SQ

Query Match 97.8%; Score 266; DB 22; Length 1500;
Best Local Similarity 98.0%; Pred. No. 1e-26;
Matches 50; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLNHEGFKLFATEATSDWLNANNVPANPVAVPWSQEGQNPSLSIRKLIRDG 51
|||||
Db 1380 QLNHEGFKLFATEATSDWLNANNVPATPVAVPWSQEGQNPSLSIRKLIRDG 1430

RESULT 6
AAU33563
ID AAU33563 standard; Protein; 344 AA.
XX
AC AAU33563;
XX
DT 14-FEB-2002 (first entry)
XX
DE Pseudomonas aeruginosa cellular proliferation protein #7.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Pseudomonas aeruginosa.
XX
PN WO200170955-A2.
XX
XX 27-SEP-2001.
PD
XX
XX 21-MAR-2001; 2001WO-US09180.
PF
XX
XX 21-MAR-2000; 2000US-191078P.
PR
XX 23-MAY-2000; 2000US-206848P.
PR
XX 26-MAY-2000; 2000US-207727P.
PR
XX 23-OCT-2000; 2000US-242578P.
PR
XX 27-NOV-2000; 2000US-253625P.
PR
XX 22-DEC-2000; 2000US-257931P.
PR
XX 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
PI
XX WPI; 2001-611495/70.
DR
XX N-PSDB; AAS51422.
DR
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
PT
XX
XX Example 3; Seq ID No 5059; 511pp; English.
PS
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets

PT cDNA contg. total human cDNA sequence
XX
PS Disclosure; Fig 2-9; 12pp; Japanese.
XX
CC This sequence is the human carbamyl phosphate synthetase I (CPSI)
CC protein. The DNA encoding this peptide or fragments of it may be
CC used as probes to detect mutations in the CPSI gene. CPSI deficiency
CC disease may be detected by using overlapping cDNA representing the
CC full length cDNA sequence of human CPSI. The DNA sequence was
CC isolated using three amplified fragments derived from protein coding
CC regions of the rat CPSI cDNA as probes.
XX
SQ Sequence 1500 AA;

Query Match 100.0%; Score 272; DB 14; Length 1500;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLHNEGFKLFATEATSDWLNANNVPANVPVAVPWSQEQNPSSLSSIRKLIRDG 51
|||||
Db 1380 QLHNEGFKLFATEATSDWLNANNVPANVPVAVPWSQEQNPSSLSSIRKLIRDG 1430

RESULT 2
AAB49222
ID AAB49222 standard; protein; 1500 AA.
XX
AC AAB49222;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human carbamyl phosphate synthase I protein.
XX
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200073322-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15079.
XX
PR 01-JUN-1999; 99US-0323472.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Summar ML, Christman BW;
XX
DE Human N1405 carbamyl phosphate synthase I protein.
XX
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200073322-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15079.
XX
PR 01-JUN-1999; 99US-0323472.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Summar ML, Christman BW;
XX
DR WPI; 2001-049936/06.
XX
CC Detecting susceptibility to suboptimal urea cycle function, e.g. bone
marrow transplant toxicity, comprises identifying a polymorphism in the
gene for carbamylphosphate synthase -
XX
PS Claim 57; Fig 12; 171pp; English.
XX
CC The present invention relates to screening for susceptibility to
suboptimal urea cycle function or to bone marrow transplant toxicity
by detecting a polymorphism in the carbamyl phosphate synthase I
(CPSI) gene. The method is used to detect subjects at risk of
hepatitis, sclerosis, pulmonary hypertension and bone marrow
transplant toxicity. These conditions can be treated or prevented
by administration of a nitric oxide precursor or by gene
therapy (administration of sequences that encode CPSI).
XX
SQ Sequence 1500 AA;

Query Match 100.0%; Score 272; DB 22; Length 1500;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLHNEGFKLFATEATSDWLNANNVPANVPVAVPWSQEQNPSSLSSIRKLIRDG 51
|||||
Db 1380 QLHNEGFKLFATEATSDWLNANNVPANVPVAVPWSQEQNPSSLSSIRKLIRDG 1430

RESULT 3
AAB49224
ID AAB49224 standard; protein; 1500 AA.
XX
AC AAB49224;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human carbamyl phosphate synthase protein.
XX
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200073322-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15079.
XX
PR 01-JUN-1999; 99US-0323472.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Summar ML, Christman BW;
XX
DE WPI; 2001-049926/06.
XX
PT Detecting susceptibility to suboptimal urea cycle function, e.g. bone
marrow transplant toxicity, comprises identifying a polymorphism in the
gene for carbamylphosphate synthase -
XX
PS Claim 57; Page 152-155; 171pp; English.
XX
CC The present invention relates to screening for susceptibility to
suboptimal urea cycle function or to bone marrow transplant toxicity
by detecting a polymorphism in the carbamyl phosphate synthase I
(CPSI) gene. The method is used to detect subjects at risk of
hepatitis, sclerosis, pulmonary hypertension and bone marrow
transplant toxicity. These conditions can be treated or prevented
by administration of a nitric oxide precursor or by gene
therapy (administration of sequences that encode CPSI).
XX
SQ Sequence 1500 AA;

Query Match 100.0%; Score 272; DB 22; Length 1500;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLHNEGFKLFATEATSDWLNANNVPANVPVAVPWSQEQNPSSLSSIRKLIRDG 51
|||||
Db 1380 QLHNEGFKLFATEATSDWLNANNVPANVPVAVPWSQEQNPSSLSSIRKLIRDG 1430

RESULT 4
AAB49225
ID AAB49225 standard; protein; 1500 AA.
XX
AC AAB49225;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human carbamyl phosphate synthase protein.
XX
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy.
XX

QY 1 QLHNEGFKLFATEATSDWLNANNVPANVPVAVPWSQEQNPSSLSSIRKLIRDG 51
|||||
Db 1380 QLHNEGFKLFATEATSDWLNANNVPANVPVAVPWSQEQNPSSLSSIRKLIRDG 1430

RESULT 3
AAB49224
ID AAB49224 standard; protein; 1500 AA.
XX
AC AAB49224;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human carbamyl phosphate synthase protein.
XX
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200073322-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15079.
XX
PR 01-JUN-1999; 99US-0323472.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Summar ML, Christman BW;
XX
DE WPI; 2001-049926/06.
XX
PT Detecting susceptibility to suboptimal urea cycle function, e.g. bone
marrow transplant toxicity, comprises identifying a polymorphism in the
gene for carbamylphosphate synthase -
XX
PS Claim 57; Page 152-155; 171pp; English.
XX
CC The present invention relates to screening for susceptibility to
suboptimal urea cycle function or to bone marrow transplant toxicity
by detecting a polymorphism in the carbamyl phosphate synthase I
(CPSI) gene. The method is used to detect subjects at risk of
hepatitis, sclerosis, pulmonary hypertension and bone marrow
transplant toxicity. These conditions can be treated or prevented
by administration of a nitric oxide precursor or by gene
therapy (administration of sequences that encode CPSI).
XX
SQ Sequence 1500 AA;

Query Match 100.0%; Score 272; DB 22; Length 1500;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLHNEGFKLFATEATSDWLNANNVPANVPVAVPWSQEQNPSSLSSIRKLIRDG 51
|||||
Db 1380 QLHNEGFKLFATEATSDWLNANNVPANVPVAVPWSQEQNPSSLSSIRKLIRDG 1430

RESULT 4
AAB49225
ID AAB49225 standard; protein; 1500 AA.
XX
AC AAB49225;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human carbamyl phosphate synthase protein.
XX
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2003, 00:29:37 ; Search time 52 Seconds
(without alignments)
155.674 Million cell updates/sec

Title: US-09-585-077C-2_COPY_1380_1430
Perfect score: 272
Sequence: 1 QLNHEGFKLFATEATSDWLN.....PSQEGQNPSLSSIRKLIRDG 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :					A_Geneseq_19Jun03:*				
Result No.	Score	Match	Length	DB ID	Description				
1	272	100.0	1500	14	AAR30636	hCPSI. Homo sapie			
2	272	100.0	1500	22	AAB49222	Human N1405 carban			
3	272	100.0	1500	22	AAB49224	Human carbamyl pho			
4	272	100.0	1500	22	AAB49225	Human carbamyl pho			
5	266	97.8	1500	22	AAB49223	Human T1405 carban			
6	67.5	24.8	344	22	AAU33563	Pseudomonas aerugi			
7	66	24.3	43	22	AAO13747	Human polypeptide			
8	64	23.5	94	22	AAO09946	Human polypeptide			
9	63.5	23.3	692	22	ABG09932	Novel human diagno			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT 1					AAR30636				
ID	XX	XX	XX	XX	AAR30636 standard; Protein: 1500 AA.				
AC	XX	XX	XX	XX	AAR30636;				
DT	XX	XX	XX	XX	06-MAY-1993 (first entry)				
DE	XX	XX	XX	XX	hCPSI.				
KW	XX	XX	XX	XX	Human; carbamyl phosphate synthetase I; CPSI; probe; mutation;				
XX	XX	XX	XX	XX	CPSI deficiency disease; rat.				
OS	XX	XX	XX	XX	Homo sapiens.				
PN	XX	XX	XX	XX	JP04335889-A.				
PD	XX	XX	XX	XX	24-NOV-1992.				
PF	XX	XX	XX	XX	09-MAY-1991; 91JP-0135902.				
PR	XX	XX	XX	XX	09-MAY-1991; 91JP-0135902.				
PA	XX	XX	XX	XX	(HARA/) HARAGUCHI Y.				
PA	XX	XX	XX	XX	(MATS/) MATSUDA I.				
XX	XX	XX	XX	XX	(MORI/) MORI M.				
DR	XX	XX	XX	XX	WPI; 1993-006237/01.				
DR	XX	XX	XX	XX	N-PSDB; AAQ34768.				
PT	XX	XX	XX	XX	Carbamyl phosphate synthetase I gene - used to detect diseases caused by carbamyl phosphatase synthetase by using overlapped				

ALIGNMENTS

```

; LENGTH: 2193
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-103660

Query Match          33.8%; Score 27.4; DB 13; Length 2193;
Best Local Similarity 59.7%; Pred. No. 1.9;
Matches 46; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 5 CATCAGACTGGCTCAACGCCCAACAATGTCCTGCGCAACCCAGTGGCATGGCCGTCCTCAAG 64
   |||| || | ||| || | ||| || | ||| || | ||| || | ||| || | ||| || |
Db 341 CATCTGAGTCTTTTCACCTCCGAATAAGCCCTCCCATCCTTGAGACCTGCAAGGTGCAAG 282

QY 65 AAGGACAGAAATCCCAAGC 81
   | | ||||| || |
Db 281 AGAGGTAGAAATCACATC 265

RESULT 14
US-10-207-655-78/c
; Sequence 78, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 78
; LENGTH: 6471
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-207-655-78

Query Match          33.8%; Score 27.4; DB 14; Length 6471;
Best Local Similarity 59.7%; Pred. No. 2.2;
Matches 46; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 5 CATCAGACTGGCTCAACGCCCAACAATGTCCTGCGCAACCCAGTGGCATGGCCGTCCTCAAG 64
   |||| || | ||| || | ||| || | ||| || | ||| || | ||| || | ||| || |
Db 332 CATCTGAGTCTTTTCACCTCCGAATAAGCCCTCCCATCCTTGAGACCTGCAAGGTGCAAG 273

QY 65 AAGGACAGAAATCCCAAGC 81
   | | ||||| || |
Db 272 AGAGGTAGAAATCACATC 256

RESULT 15
US-10-068-674-1/c
; Sequence 1, Application US/10068674
; Publication No. US20020177203A1
; GENERAL INFORMATION:
; APPLICANT: Kevirikko, Kari
; APPLICANT: Pihlajaniemi, Taina A.
; TITLE OF INVENTION: 2 SUBUNIT OF PROLYL-4-HYDROXYLASE,
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE ENCODING SUCH SUBUNIT AND METHODS
; TITLE OF INVENTION: FOR PRODUCING THE SAME
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/068,674
; FILING DATE: 06-FEB-2002
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/633,879
; FILING DATE: 10-MAR-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 8389-041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3694
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2168 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 151..1761
US-10-068-674-1

Query Match          32.6%; Score 26.4; DB 13; Length 2168;
Best Local Similarity 61.8%; Pred. No. 4.4;
Matches 42; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 9 AGACTGGCTCAACGCCCAACAATGTCCTGCGCAACCCAGTGGCATGGCCGTCCTCAAGAAGG 68
   || ||||| || ||| || | ||| || | |||| || | |||| || | ||| ||
Db 2005 AGTCTGGCTCTACCCGACGCGCATGCCAAGACACCCCGACAGAGCCAGGCGTGAAGCTGG 1946

QY 69 ACAGAATC 76
   |||| || |
Db 1945 TCAGAGTC 1938
```

Search completed: September 13, 2003, 02:10:57
Job time : 166 secs

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCGCCAAACCCAGTGGCATGGCCGCTCT 60
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCGCCAAACCCAGTGGCATGGCCGCTCT 216

QY 61 CAAGAAGGACAGAAATCCCAGC 81
Db 217 CAAGAAGGACAGAAATCCCAGC 237

RESULT 6
US-10-017-754-93
; Sequence 93, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-93

Query Match 98.0%; Score 79.4; DB 14; Length 531;
Best Local Similarity 98.8%; Pred. No. 6.3e-20;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCGCCAAACCCAGTGGCATGGCCGCTCT 60
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCGCCAAACCCAGTGGCATGGCCGCTCT 216

QY 61 CAAGAAGGACAGAAATCCCAGC 81
Db 217 CAAGAAGGACAGAAATCCCAGC 237

RESULT 7
US-10-101-510-553
; Sequence 553, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 553
; LENGTH: 5772
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (5240)..(5365)

; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-101-510-553

Query Match 98.0%; Score 79.4; DB 12; Length 5772;
Best Local Similarity 98.8%; Pred. No. 8.8e-20;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCGCCAAACCCAGTGGCATGGCCGCTCT 60
Db 4310 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCGCCAAACCCAGTGGCATGGCCGCTCT 4369

QY 61 CAAGAAGGACAGAAATCCCAGC 81
Db 4370 CAAGAAGGACAGAAATCCCAGC 4390

RESULT 8
US-09-960-352-12478
; Sequence 12478, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12478
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 53-LIB34-037-Q1-E1-F2
US-09-960-352-12478

Query Match 90.1%; Score 73; DB 10; Length 434;
Best Local Similarity 93.8%; Pred. No. 1.5e-17;
Matches 76; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCGCCAAACCCAGTGGCATGGCCGCTCT 60
Db 36 GCCACATCAGACTGGCTCAATGCCAACAATGTACTCTACCTACCCAGTGGCATGGCCATCT 95

QY 61 CAAGAAGGACAGAAATCCCAGC 81
Db 96 CAAGAAGGACAGAAATCCCAGC 116

RESULT 9
US-09-960-352-1818
; Sequence 1818, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 1818
; LENGTH: 242
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 08-LIB34-052-Q1-E1-B11
US-09-960-352-1818

```
RESULT 2
US-09-736-457-93
; Sequence 93, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736/457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-93

Query Match      98.0%; Score 79.4; DB 10; Length 531;
Best Local Similarity 98.8%; Pred. No. 6.3e-20;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCAACCCAGTGGCGGTCT 60
    |||||
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCAACCCAGTGGCGGTCT 216

QY 61 CAAGAAGGACAGATCCCAGC 81
    |||||
Db 217 CAAGAAGGACAGATCCCAGC 237

RESULT 3
US-09-902-941-93
; Sequence 93, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-93

Query Match      98.0%; Score 79.4; DB 10; Length 531;
Best Local Similarity 98.8%; Pred. No. 6.3e-20;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCAACCCAGTGGCGGTCT 60
    |||||
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCAACCCAGTGGCGGTCT 216

QY 61 CAAGAAGGACAGATCCCAGC 81
    |||||
Db 217 CAAGAAGGACAGATCCCAGC 237

RESULT 4
US-09-849-626-93
; Sequence 93, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-849-626-93

Query Match      98.0%; Score 79.4; DB 10; Length 531;
Best Local Similarity 98.8%; Pred. No. 6.3e-20;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCAACCCAGTGGCGGTCT 60
    |||||
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCAACCCAGTGGCGGTCT 216

QY 61 CAAGAAGGACAGATCCCAGC 81
    |||||
Db 217 CAAGAAGGACAGATCCCAGC 237

RESULT 5
US-09-476-300-93
; Sequence 93, Application US/09476300
; Publication No. US20030125245A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C3
; CURRENT APPLICATION NUMBER: US/09/476,300
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 785
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-476-300-93

Query Match      98.0%; Score 79.4; DB 11; Length 531;
Best Local Similarity 98.8%; Pred. No. 6.3e-20;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 13, 2003, 00:27:37 ; Search time 151 Seconds
(without alignments)
1302.306 Million cell updates/sec

Title: US-09-585-077c-1_COPY_4300_4380
Perfect score: 81
Sequence: 1 gccacatcagactggtctcaa.....aagaaggacagaatcccagc 81

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1632420 seqs, 1213878141 residues

Total number of hits satisfying chosen parameters: 3264840

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published_Applications_NA:*
- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
 - 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
 - 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
 - 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
 - 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
 - 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
 - 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
 - 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	5215	10	US-09-880-107-1769
2	79.4	98.0	531	10	US-09-736-457-93
3	79.4	98.0	531	10	US-09-902-941-93
4	79.4	98.0	531	10	US-09-849-626-93
5	79.4	98.0	531	11	US-09-476-300-93
6	79.4	98.0	531	14	US-10-017-754-93
7	79.4	98.0	5772	12	US-10-101-510-553
8	73	90.1	434	10	US-09-950-352-12478
9	40.2	49.6	242	10	US-09-960-352-1818
10	30.6	37.8	31	9	US-09-801-274-1768
11	29.4	36.3	3591	11	US-09-764-891-5507
c 12	27.4	33.8	2193	13	US-10-027-97735
c 13	27.4	33.8	2193	13	US-10-027-632-103660
c 14	27.4	33.8	6471	14	US-10-207-655-78
c 15	26.4	32.6	2168	13	US-10-068-674-1
16	26	32.1	375	14	US-10-106-698-3792
					Sequence 1769, Ap
					Sequence 93, Appl
					Sequence 93, Appl
					Sequence 93, Appl
					Sequence 93, Appl
					Sequence 93, Appl
					Sequence 553, App
					Sequence 12478, A
					Sequence 1818, Ap
					Sequence 1768, Ap
					Sequence 5507, Ap
					Sequence 97735, A
					Sequence 103660,
					Sequence 78, Appl
					Sequence 1, Appli
					Sequence 3792, Ap

c 17	26	32.1	987	10	US-09-738-973-289	Sequence 289, App
c 18	26	32.1	987	10	US-09-854-133-289	Sequence 289, App
c 19	26	32.1	987	14	US-10-144-649A-289	Sequence 289, App
20	25.6	31.6	2985	10	US-09-947-953-1	Sequence 1, Appli
21	25.4	31.4	288	10	US-09-983-965-4858	Sequence 4858, Ap
22	25.4	31.4	26225	9	US-09-764-869-1276	Sequence 1276, Ap
23	25.4	31.4	26225	14	US-10-091-504-1276	Sequence 1276, Ap
24	25	30.9	331	10	US-09-783-590-7357	Sequence 7357, Ap
25	25	30.9	414	10	US-09-833-381-1254	Sequence 1254, Ap
c 26	24.8	30.6	284	10	US-09-833-381-417	Sequence 417, App
27	24.8	30.6	957	13	US-10-027-632-163409	Sequence 163409,
c 28	24.8	30.6	969	10	US-09-280-197-13	Sequence 13, Appl
c 29	24.8	30.6	3276	10	US-09-280-197-4	Sequence 4, Appli
c 30	24.8	30.6	3276	10	US-09-423-126-8	Sequence 8, Appli
c 31	24.8	30.6	4851	10	US-09-712-363-116	Sequence 116, App
c 32	24.6	30.4	723	13	US-10-027-632-165578	Sequence 165578,
c 33	24.4	30.1	1702	10	US-09-880-107-3680	Sequence 3680, Ap
c 34	24.2	29.9	576	13	US-10-027-632-45850	Sequence 45850, A
c 35	24.2	29.9	576	13	US-10-027-632-61618	Sequence 61618, A
c 36	24.2	29.9	576	13	US-10-027-632-61619	Sequence 61619, A
c 37	24.2	29.9	576	13	US-10-027-632-107709	Sequence 107709,
c 38	24.2	29.9	576	13	US-10-027-632-107710	Sequence 107710,
39	24.2	29.9	681	13	US-10-027-632-142706	Sequence 142706,
40	24.2	29.9	681	13	US-10-027-632-142707	Sequence 142707,
c 41	24.2	29.9	1281	14	US-10-156-761-2272	Sequence 2272, Ap
c 42	24.2	29.9	9025608	14	US-10-156-761-1	Sequence 1, Appli
43	24	29.6	469	11	US-09-918-995-2365	Sequence 2365, Ap
44	24	29.6	870	13	US-10-027-632-159815	Sequence 159815,
c 45	24	29.6	1263	10	US-09-738-626-2281	Sequence 2281, Ap

ALIGNMENTS

RESULT 1
US-09-880-107-1769
; Sequence 1769, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1769
; LENGTH: 5215
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D90282
US-09-880-107-1769

Query Match	100.0%;	Score 81;	DB 10;	Length 5215;
Best Local Similarity	100.0%;	Pred. No. 2.2e-20;		
Matches	81;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCAACCCAGTGGCATGGCCGCTCT	60	
Db	4295	GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCAACCCAGTGGCATGGCCGCTCT	4354	
QY	61	CAAGAAGGACAGATCCCAGC	81	
Db	4355	CAAGAAGGACAGATCCCAGC	4375	

Search completed: September 13, 2003, 00:29:22
Job time : 64.5 secs


```

Query Match          31.6%; Score 25.6; DB 4; Length 1036;
Best Local Similarity 57.5%; Pred. No. 6.4;
Matches 46; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY      2 CCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCAGTGGCCGCTTC 61
      ||| | | ||| || |||| || ||| || ||||| | | | | |
Db      6 CCCCCACCTCCAGGGCCACCGCGACTTTGTACCCCCCAACCCCTGAGGAAGATGGGGGC 65

QY      62 AAGAAGGACAGAAATCCGAGC 81
      ||||| | | |||| | |
Db      66 AAGAAGATCACGCTCCCCGC 85

RESULT 14
US-09-252-991A-12606/c
; Sequence 12606, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```

```

; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match          32.3%; Score 26.2; DB 3; Length 4403765;
Best Local Similarity 67.3%; Pred. No. 27;
Matches 37; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 18 CAACGCCAACAAATGTCCCTGCCAACCCAGTGGCATGGCCGTCTCAAGAAGGACAG 72
Db 3253074 CACGCCACCGTTGCCCGCGCGAGCCGTTAGCACCCGCTGGATCACGAAGGACAG 3253128
      || ||||| | || || ||||| || ||||| || || ||||| |||||
      || ||||| | || || ||||| || ||||| || || ||||| |||||

RESULT 10
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match          32.3%; Score 26.2; DB 3; Length 4411529;
Best Local Similarity 67.3%; Pred. No. 27;
Matches 37; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 18 CAACGCCAACAAATGTCCCTGCCAACCCAGTGGCATGGCCGTCTCAAGAAGGACAG 72
Db 3258798 CACCGCACCGTTGCCCGCGCGAGCCGTTAGCACCCGCTGGATCACGAAGGACAG 3258852
      || ||||| | || || ||||| || ||||| || || ||||| |||||
      || ||||| | || || ||||| || ||||| || || ||||| |||||

RESULT 11
US-09-370-838-289/c
; Sequence 289, Application US/09370838
; Patent No. 644425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 289

```



```

; SEQ ID NO 11
; LENGTH: 5761
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)..(4626)
US-09-323-472A-11

Query Match      100.0%; Score 81; DB 4; Length 5761;
Best Local Similarity 100.0%; Pred. No. 6.6e-19;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCCAACCCAGTGGCATGGCCGTCT 60
Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCCAACCCAGTGGCATGGCCGTCT 4359

QY 61 CAAGAAGGACAGAAATCCCAGC 81
Db 4360 CAAGAAGGACAGAAATCCCAGC 4380

RESULT 3
US-09-323-472A-13
; Sequence 13, Application US/09323472A
; Patent No. 6346382
; GENERAL INFORMATION:
; APPLICANT: Summar, Marshall
; APPLICANT: Christman, Brian
; TITLE OF INVENTION: HUMAN CARBAMYL PHOSPHATE SYNTHETASE I POLYMORPHISM AND DIAGNOSTIC
; TITLE OF INVENTION: THERAPEUTIC METHODS RELATING THERETO
; FILE REFERENCE: 1242/19
; CURRENT APPLICATION NUMBER: US/09/323,472A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 5762
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)..(4626)
US-09-323-472A-13

Query Match      100.0%; Score 81; DB 4; Length 5762;
Best Local Similarity 100.0%; Pred. No. 6.6e-19;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCCAACCCAGTGGCATGGCCGTCT 60
Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCCAACCCAGTGGCATGGCCGTCT 4359

QY 61 CAAGAAGGACAGAAATCCCAGC 81
Db 4360 CAAGAAGGACAGAAATCCCAGC 4380

RESULT 4
US-09-323-472A-5
; Sequence 5, Application US/09323472A
; Patent No. 6346382
; GENERAL INFORMATION:
; APPLICANT: Summar, Marshall
; APPLICANT: Christman, Brian
; TITLE OF INVENTION: HUMAN CARBAMYL PHOSPHATE SYNTHETASE I POLYMORPHISM AND DIAGNOSTIC
; TITLE OF INVENTION: THERAPEUTIC METHODS RELATING THERETO
; FILE REFERENCE: 1242/19
; CURRENT APPLICATION NUMBER: US/09/323,472A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
```

```

; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (111)..(224)
; OTHER INFORMATION: n is G or A or C or T/U
US-09-323-472A-5

Query Match      98.0%; Score 79.4; DB 4; Length 495;
Best Local Similarity 98.8%; Pred. No. 1.3e-18;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCCAACCCAGTGGCATGGCCGTCT 60
Db 126 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCCAACCCAGTGGCATGGCCGTCT 185

QY 61 CAAGAAGGACAGAAATCCCAGC 81
Db 186 CAAGAAGGACAGAAATCCCAGC 206

RESULT 5
US-09-702-705-93
; Sequence 93, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-93

Query Match      98.0%; Score 79.4; DB 4; Length 531;
Best Local Similarity 98.8%; Pred. No. 1.3e-18;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCCAACCCAGTGGCATGGCCGTCT 60
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCCAACCCAGTGGCATGGCCGTCT 216

QY 61 CAAGAAGGACAGAAATCCCAGC 81
Db 217 CAAGAAGGACAGAAATCCCAGC 237

RESULT 6
US-09-736-457-93
; Sequence 93, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 22:27:31 ; Search time 44.5 Seconds
(without alignments)
803.416 Million cell updates/sec

Title: US-09-585-077C-1_COPY_4300_4380
Perfect score: 81
Sequence: 1 gccacatcagactggctcaa.....aagaaggacagaatcccgac 81

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	5761	4	US-09-323-472A-1
2	81	100.0	5761	4	US-09-323-472A-11
3	81	100.0	5762	4	US-09-323-472A-13
4	79.4	98.0	495	4	US-09-323-472A-5
5	79.4	98.0	531	4	US-09-702-705-93
6	79.4	98.0	531	4	US-09-736-457-93
7	79.4	98.0	5761	4	US-09-323-472A-3
8	26.4	32.6	2168	2	US-08-633-879C-1
9	26.2	32.3	4403765	3	US-09-103-840A-2
10	26.2	32.3	4411529	3	US-09-103-840A-1
11	26	32.1	987	4	US-09-370-838-289
12	25.6	31.6	468	3	US-09-067-782A-6
13	25.6	31.6	1036	4	US-09-016-434-1308
14	24.8	30.6	654	4	US-09-252-991A-12606
15	24.8	30.6	3276	3	US-08-633-768A-4
16	24	29.6	432	4	US-09-252-991A-16133
17	24	29.6	438	4	US-09-252-991A-8292
18	23.8	29.4	1703	3	US-09-135-021-77
19	23.8	29.4	1703	3	US-09-135-020-3
20	23.8	29.4	1703	3	US-09-135-010A-3
21	23.8	29.4	1703	4	US-09-444-871-3
22	23.8	29.4	1703	4	US-09-597-735-3
23	23.8	29.4	1703	4	US-09-444-295-3
24	23.8	29.4	1703	4	US-09-597-732-3
25	23.8	29.4	1703	4	US-09-597-731-3
26	23.6	29.1	1056	3	US-09-067-782A-1
27	23.6	29.1	1520	5	PCT-US93-07213-8

c	28	23.6	29.1	1820	2	US-08-828-922-2	Sequence 2, Appli
c	29	23.6	29.1	2219	5	PCT-US93-07213-4	Sequence 4, Appli
c	30	23.6	29.1	2870	5	PCT-US93-07213-3	Sequence 3, Appli
c	31	23.6	29.1	3102	5	PCT-US93-07213-1	Sequence 1, Appli
c	32	23.6	29.1	152331	3	US-09-128-155-16	Sequence 16, Appl
c	33	23.4	28.9	519	4	US-09-844-634-19	Sequence 19, Appl
c	34	23.4	28.9	1017	4	US-09-252-991A-11573	Sequence 11573, A
c	35	23.4	28.9	1056	4	US-09-252-991A-11806	Sequence 11806, A
c	36	23.4	28.9	1140	4	US-09-252-991A-11515	Sequence 11515, A
c	37	23.4	28.9	3683	4	US-09-844-634-3	Sequence 3, Appli
c	38	23.4	28.9	15602	4	US-09-844-634-17	Sequence 17, Appl
c	39	23.2	28.6	1900	4	US-09-608-285A-47	Sequence 47, Appl
c	40	23.2	28.6	1900	4	US-09-557-800C-47	Sequence 47, Appl
c	41	23.2	28.6	2294	4	US-09-608-285A-49	Sequence 49, Appl
c	42	23.2	28.6	2294	4	US-09-557-800C-49	Sequence 49, Appl
c	43	23.2	28.6	2371	4	US-09-608-285A-46	Sequence 46, Appl
c	44	23.2	28.6	2371	4	US-09-557-800C-46	Sequence 46, Appl
c	45	23.2	28.6	2497	4	US-09-608-285A-51	Sequence 51, Appl

ALIGNMENTS

RESULT 1
US-09-323-472A-1
; Sequence 1, Application US/09323472A
; Patent No. 6346382
; GENERAL INFORMATION:
; APPLICANT: Summar, Marshall
; APPLICANT: Christman, Brian
; TITLE OF INVENTION: HUMAN CARBAMYL PHOSPHATE SYNTHETASE I POLYMORPHISM AND DIAGN
; TITLE OF INVENTION: THERAPEUTIC METHODS RELATING THERETO
; FILE REFERENCE: 1242/19
; CURRENT APPLICATION NUMBER: US/09/323,472A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 5761
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)..(4626)
US-09-323-472A-1

Query Match	100.0%;	Score 81;	DB 4;	Length 5761;
Best Local Similarity	100.0%;	Pred. No. 6.6e-19;		
Matches	81;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCAACCCAGTGGCGTCT 60		
Db	4300	GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCAACCCAGTGGCGTCT 4359		
Qy	61	CAAGAAGGACAGATCCCAGC 81		
Db	4360	CAAGAAGGACAGATCCCAGC 4380		

RESULT 2
US-09-323-472A-11
; Sequence 11, Application US/09323472A
; Patent No. 6346382
; GENERAL INFORMATION:
; APPLICANT: Summar, Marshall
; APPLICANT: Christman, Brian
; TITLE OF INVENTION: HUMAN CARBAMYL PHOSPHATE SYNTHETASE I POLYMORPHISM AND DIAGN
; TITLE OF INVENTION: THERAPEUTIC METHODS RELATING THERETO
; FILE REFERENCE: 1242/19
; CURRENT APPLICATION NUMBER: US/09/323,472A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0

mRNA sequence.
BG533520
BG533520.1 GI:13525060
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 949)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM918 row: f column: 09
High quality sequence stop: 740.
FEATURES
source Location/Qualifiers
1..949
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4071944"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_76"
/note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgctggcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGCGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGGCGCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.85 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."
BASE COUNT 271 a 224 c 181 g 272 t 1 others
ORIGIN
Query Match 98.0%; Score 79.4; DB 10; Length 949;
Best Local Similarity 98.8%; Pred. No. 8.6e-15;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 60
|||||
Db 181 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCT 240
QY 61 CAAGAAGGACAGAAATCCCAGC 81
|||||
Db 241 CAAGAAGGACAGAAATCCCAGC 261
RESULT 15
AV656423
LOCUS AV656423 362 bp mRNA linear EST 16-JAN-2002
DEFINITION AV656423 GLC Homo sapiens cDNA clone GLCERB10 3', mRNA sequence.
ACCESSION AV656423
VERSION AV656423.1 GI:9877437
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 362)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z. and Han,Z.

TITLE Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE 21625106
PUBMED 11752456
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source Location/Qualifiers
1..362
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCERB10"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 106 a 92 c 71 g 93 t
ORIGIN
Query Match 96.8%; Score 78.4; DB 9; Length 362;
Best Local Similarity 98.8%; Pred. No. 1.2e-14;
Matches 79; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 CCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCTC 61
|||||
Db 1 CCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCACCCAGTGGCATGGCCGTCTC 60
QY 62 AAGAAGGACAGAAATCCCAGC 81
|||||
Db 61 AAGAAGGACAGAAATCCCAGC 80
Search completed: September 13, 2003, 00:27:26
Job time : 1968 secs

FEATURES source Location/Qualifiers 1..636 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="GLCDYC09" /tissue_type="corresponding non cancerous liver tissue" /dev_stage="Adult" /lab_host="SOLR" /clone_lib="GLC" /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI" BASE COUNT 179 a 148 c 122 g 185 t 2 others

Query Match 98.0%; Score 79.4; DB 9; Length 636; Best Local Similarity 98.8%; Pred. No. 7.4e-15; Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCGGTCT 60 ||||| |

Db 59 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCGGTCT 118 ||||| |

QY 61 CAAGAAGGACAGATCCCGAGC 81 ||||| |

Db 119 CAAGAAGGACAGATCCCGAGC 139 ||||| |

RESULT 12 AV683932 LOCUS AV683932 GKc Homo sapiens cDNA clone GKCDID08 5', mRNA linear EST 16-JAN-2002 DEFINITION AV683932 ACCESSION AV683932 VERSION AV683932.1 GI:10285795 KEYWORDS EST. SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 637) XIAO,H., QU,J., LIU,F., HUANG,Q., CHENG,Z., LI,N., DU,J., HU,W., SHEN,K., LU,G., FU,G., ZHONG,M., XU,S., GU,W., HUANG,W., ZHAO,X., HU,G., GU,J., CHEN,Z. and HAN,Z. Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001) 21625106 PUBMED 11752456 Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai. Location/Qualifiers 1..637 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="GKCDID08" /tissue_type="hepatocellular carcinoma" /dev_stage="Adult" /lab_host="SOLR" /clone_lib="GKc" /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI" BASE COUNT 173 a 147 c 118 g 196 t 3 others

FEATURES source Location/Qualifiers 1..636 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="DKFZp434E087" /tissue_type="testis" /dev_stage="adult" /lab_host="DH10B" /clone_lib="434 (synonym: htes3)" /note="Vector: pSport1; Site_1: NotI; Site_2: SalI" BASE COUNT 171 a 128 c 156 g 190 t 1 others

Query Match 98.0%; Score 79.4; DB 9; Length 646; Best Local Similarity 98.8%; Pred. No. 7.4e-15; Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCGGTCT 60 ||||| |

Db 469 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCGGTCT 410 ||||| |

QY 61 CAAGAAGGACAGATCCCGAGC 81 ||||| |

Db 409 CAAGAAGGACAGATCCCGAGC 389 ||||| |

RESULT 14 BG533520 LOCUS BG533520 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4071944 5', DEFINITION 601860935F2

FEATURES source Location/Qualifiers 1..646 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="DKFZp434E087" /tissue_type="testis" /dev_stage="adult" /lab_host="DH10B" /clone_lib="434 (synonym: htes3)" /note="Vector: pSport1; Site_1: NotI; Site_2: SalI" BASE COUNT 171 a 128 c 156 g 190 t 1 others

Query Match 98.0%; Score 79.4; DB 9; Length 646; Best Local Similarity 98.8%; Pred. No. 7.4e-15; Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCGGTCT 60 ||||| |

Db 469 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCGGTCT 410 ||||| |

QY 61 CAAGAAGGACAGATCCCGAGC 81 ||||| |

Db 409 CAAGAAGGACAGATCCCGAGC 389 ||||| |

RESULT 14 BG533520 LOCUS BG533520 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4071944 5', DEFINITION 601860935F2

FEATURES source Location/Qualifiers 1..636 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="GLCDYC09" /tissue_type="corresponding non cancerous liver tissue" /dev_stage="Adult" /lab_host="SOLR" /clone_lib="GLC" /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI" BASE COUNT 179 a 148 c 122 g 185 t 2 others

Query Match 98.0%; Score 79.4; DB 9; Length 636; Best Local Similarity 98.8%; Pred. No. 7.4e-15; Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCGGTCT 60 ||||| |

Db 59 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCGGTCT 118 ||||| |

QY 61 CAAGAAGGACAGATCCCGAGC 81 ||||| |

Db 119 CAAGAAGGACAGATCCCGAGC 139 ||||| |

RESULT 12 AV683932 LOCUS AV683932 GKc Homo sapiens cDNA clone GKCDID08 5', mRNA linear EST 16-JAN-2002 DEFINITION AV683932 ACCESSION AV683932 VERSION AV683932.1 GI:10285795 KEYWORDS EST. SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 637) XIAO,H., QU,J., LIU,F., HUANG,Q., CHENG,Z., LI,N., DU,J., HU,W., SHEN,K., LU,G., FU,G., ZHONG,M., XU,S., GU,W., HUANG,W., ZHAO,X., HU,G., GU,J., CHEN,Z. and HAN,Z. Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001) 21625106 PUBMED 11752456 Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai. Location/Qualifiers 1..637 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="GKCDID08" /tissue_type="hepatocellular carcinoma" /dev_stage="Adult" /lab_host="SOLR" /clone_lib="GKc" /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI" BASE COUNT 173 a 147 c 118 g 196 t 3 others

FEATURES source Location/Qualifiers 1..636 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="DKFZp434E087" /tissue_type="testis" /dev_stage="adult" /lab_host="DH10B" /clone_lib="434 (synonym: htes3)" /note="Vector: pSport1; Site_1: NotI; Site_2: SalI" BASE COUNT 171 a 128 c 156 g 190 t 1 others

Query Match 98.0%; Score 79.4; DB 9; Length 646; Best Local Similarity 98.8%; Pred. No. 7.4e-15; Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCGGTCT 60 ||||| |

Db 469 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCGGTCT 410 ||||| |

QY 61 CAAGAAGGACAGATCCCGAGC 81 ||||| |

Db 409 CAAGAAGGACAGATCCCGAGC 389 ||||| |

RESULT 14 BG533520 LOCUS BG533520 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4071944 5', DEFINITION 601860935F2

FEATURES source Location/Qualifiers 1..646 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="DKFZp434E087" /tissue_type="testis" /dev_stage="adult" /lab_host="DH10B" /clone_lib="434 (synonym: htes3)" /note="Vector: pSport1; Site_1: NotI; Site_2: SalI" BASE COUNT 171 a 128 c 156 g 190 t 1 others

Query Match 98.0%; Score 79.4; DB 9; Length 646; Best Local Similarity 98.8%; Pred. No. 7.4e-15; Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCGGTCT 60 ||||| |

Db 469 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCGGTCT 410 ||||| |

QY 61 CAAGAAGGACAGATCCCGAGC 81 ||||| |

Db 409 CAAGAAGGACAGATCCCGAGC 389 ||||| |

RESULT 14 BG533520 LOCUS BG533520 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4071944 5', DEFINITION 601860935F2

RESULT 9
AV661502
LOCUS AV661502 564 bp mRNA linear EST 16-JAN-2002
DEFINITION AV661502 GLC Homo sapiens cDNA clone GLCGTB04 3', mRNA sequence.
ACCESSION AV661502
VERSION AV661502.1 GI:9882516
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 564)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE 21625106
PUBMED 11752456
COMMENT Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
Location/Qualifiers
1..564
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCGTB04"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOIR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 140 a 140 c 129 g 155 t
ORIGIN
Query Match 98.0%; Score 79.4; DB 9; Length 564;
Best Local Similarity 98.8%; Pred. No. 7e-15;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAAGCCCAACAATGTCCCTGCCAACCCAGTGGCATGGCGGTCT 60
|||||
Db 453 GCCACATCAGATGGCTCAAGCCCAACAATGTCCCTGCCAACCCAGTGGCATGGCGGTCT 512
QY 61 CAGAAGGACAGATCCCAGC 81
|||||
Db 513 CAGAAGGACAGATCCCAGC 533
RESULT 10
CB153827
LOCUS CB153827 625 bp mRNA linear EST 29-JAN-2003
DEFINITION K-EST0211442 L17N670205 Homo sapiens cDNA clone L17N670205-7-B10
5', mRNA sequence.
ACCESSION CB153827
VERSION CB153827.1 GI:28138823
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 625)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and

Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 7 row: B column: 10
High quality sequence stop: 625.
FEATURES
source
Location/Qualifiers
1..625
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L17N670205-7-B10"
/sex="F"
/lab_host="Top10F"
/clone_lib="L17N670205"
/note="Organ: Liver; Vector: pT7T3-Pac; Site_1: EcoRI;
Site_2: NotI; The library was contributed by the Soares
laboratory and it was constructed as described by Bonalao,
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
6(9); 791-806. RNA was prepared from harvested cell
culture."
BASE COUNT 178 a 159 c 134 g 154 t
ORIGIN
Query Match 98.0%; Score 79.4; DB 14; Length 625;
Best Local Similarity 98.8%; Pred. No. 7.3e-15;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAAGCCCAACAATGTCCCTGCCAACCCAGTGGCATGGCGGTCT 60
|||||
Db 221 GCCACATCAGACTGGCTCAAGCCCAACAATGTCCCTGCCAACCCAGTGGCATGGCGGTCT 280
QY 61 CAGAAGGACAGATCCCAGC 81
|||||
Db 281 CAGAAGGACAGATCCCAGC 301
RESULT 11
AV654671
LOCUS AV654671 636 bp mRNA linear EST 15-JAN-2002
DEFINITION AV654671 GLC Homo sapiens cDNA clone GLCDBC09 3', mRNA sequence.
ACCESSION AV654671
VERSION AV654671.1 GI:9875685
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 636)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE 21625106
PUBMED 11752456
COMMENT Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

source
Location/Qualifiers
1. .374
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCGNF11"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 91 a 103 c 87 g 93 t
ORIGIN
Query Match 98.0%; Score 79.4; DB 9; Length 374;
Best Local Similarity 98.8%; Pred. No. 6e-15;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
Db 277 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 336
QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 337 CAAGAAGGACAGAATCCCAGC 357

RESULT 7

AI174790
LOCUS HA2511 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
DEFINITION
sequence.
ACCESSION AI174790
VERSION AI174790.1 GI:6361174
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 501)
AUTHORS Yu.Y., Zhang.C., Luo.L., Ouyang.S., Li.W., Wu.J., Zhou.S., Liu.M.
and He,F.
TITLE Expression profile analysis of a human fetal liver cDNA library
JOURNAL Unpublished
COMMENT Contact: Chenggang Zhang
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Email: zhang.chenggang@hotmail.com.

FEATURES

source
Location/Qualifiers
1. .501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="MC1061/p3"
/clone_lib="Human fetal liver cDNA library"
/note="Vector: pCDNA1"
BASE COUNT 143 a 124 c 110 g 124 t
ORIGIN

Query Match 98.0%; Score 79.4; DB 9; Length 501;
Best Local Similarity 98.8%; Pred. No. 6.7e-15;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
Db 181 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 240

QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 241 CAAGAAGGACAGAATCCCAGC 261

RESULT 8

BM821786
LOCUS BM821786
DEFINITION K-EST091028 S20T665307 Homo sapiens cDNA clone S20T665307-14-F07
5', mRNA sequence.
ACCESSION BM821786
VERSION BM821786.1 GI:19178199
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 540)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished
COMMENT Contact: Kim YS
Genome Research Center

Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 14 row: F column: 07
High quality sequence stop: 540.
Location/Qualifiers
1. .540
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S20T665307-14-F07"
/sex="M"
/lab_host="Top10F"
/clone_lib="S20T665307"

/note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
BASE COUNT 147 a 133 c 121 g 139 t
ORIGIN

Query Match 98.0%; Score 79.4; DB 12; Length 540;
Best Local Similarity 98.8%; Pred. No. 6.9e-15;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
Db 329 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 388

QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 389 CAAGAAGGACAGAATCCCAGC 409

Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCARCCAGTGGCATGGCCGTCT 60
|||||

Db 130 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCARCCAGTGGCATGGCCGTCT 189
|||||

QY 61 CAAGAAGGACAGATCCCGC 81
|||||

Db 190 CAAGAAGGACAGATCCCGC 210
|||||

RESULT 4
AI065054
LOCUS
DEFINITION HA0887 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
sequence.
ACCESSION AI065054
VERSION AI065054.1 GI:6359326
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 490)
AUTHORS Yu,Y., Zhang,C., Luo,L., Ouyang,S., Li,W., Wu,J., Zhou,S., Liu,M.
and He,F.
TITLE Expression profile analysis of a human fetal liver cDNA library
JOURNAL Unpublished
COMMENT Contact: Yongtao Yu
Department of Hematology
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Tel: 0086-10-68159479
Fax: 0086-10-68214653
Email: yyt48@yahoo.com.
FEATURES
source
1. 490
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="MC1061/P3"
/clone_lib="Human fetal liver cDNA library"
/note="Vector: pCDNA1"
BASE COUNT 134 a. 124 c 105 g 123 t 4 others
ORIGIN
Query Match 98.8%; Score 80; DB 9; Length 490;
Best Local Similarity 98.8%; Pred. No. 4.3e-15;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
|||||

Db 212 GCCACATCANACTGGCTCAACGCCAACAAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 271
|||||

QY 61 CAAGAAGGACAGATCCCGC 81
|||||

Db 272 CAAGAAGGACAGATCCCGC 292
|||||

RESULT 5
AV696724
LOCUS
DEFINITION AV696724 GKC Homo sapiens cDNA clone GKCBUC07 5', mRNA sequence.
ACCESSION AV696724
VERSION AV696724.1 GI:10298587
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 366)

AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
JOURNAL 21625106
MEDLINE 11752456
PUBMED
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1. 366
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GKCBUC07"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GKC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 92 a 98 c 82 g 93 t 1 others
ORIGIN
Query Match 98.0%; Score 79.4; DB 9; Length 366;
Best Local Similarity 98.8%; Pred. No. 6e-15;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
|||||

Db 193 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 252
|||||

QY 61 CAAGAAGGACAGATCCCGC 81
|||||

Db 253 CAAGAAGGACAGATCCCGC 273
|||||

RESULT 6
AV661012
LOCUS
DEFINITION AV661012 GLC Homo sapiens cDNA clone GLCGNfl1 3', mRNA sequence.
ACCESSION AV661012
VERSION AV661012.1 GI:9882026
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 374)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
JOURNAL 21625106
MEDLINE 11752456
PUBMED
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)

```
FEATURES
source
1. .699
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4858442"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_49"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT      185 a      166 c      158 g      190 t
ORIGIN

Query Match      100.0%; Score 81; DB 12; Length 699;
Best Local Similarity 100.0%; Pred. No. 2.4e-15;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCCTGCT 60
|||||
Db 517 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCCTGCT 576
|||||

QY 61 CAAGAAGGACAGATCCCAGC 81
|||||
Db 577 CAAGAAGGACAGATCCCAGC 597
|||||

RESULT 2
BG616938
LOCUS
DEFINITION
602615612F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4734387 5',
mRNA sequence.
ACCESSION
BG616938
VERSION
BG616938.1 GI:13668309
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 706)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1597 row: h column: 04
High quality sequence stop: 706.
Location/Qualifiers
1. .706
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4734387"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_76"
/note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgctcgcc); Site_2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGGCGGCGCATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
BASE COUNT      256 a      183 c      181 g      223 t
ORIGIN

Query Match      100.0%; Score 81; DB 10; Length 843;
Best Local Similarity 100.0%; Pred. No. 2.6e-15;
```

```
5'-ATTCTAGAGCGGCGGCGGCGCATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.85
kb (range 1.0-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
BASE COUNT      200 a      173 c      151 g      181 t      1 others
ORIGIN

Query Match      100.0%; Score 81; DB 10; Length 706;
Best Local Similarity 100.0%; Pred. No. 2.4e-15;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCCTGCT 60
|||||
Db 283 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCCTGCT 342
|||||

QY 61 CAAGAAGGACAGATCCCAGC 81
|||||
Db 343 CAAGAAGGACAGATCCCAGC 363
|||||

RESULT 3
BE971350
LOCUS
DEFINITION
601651514F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:3934971 5',
mRNA sequence.
ACCESSION
BE971350
VERSION
BE971350.1 GI:10584686
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 843)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM776 row: k column: 04
High quality sequence stop: 583.
Location/Qualifiers
1. .843
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3934971"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_81"
/note="Organ: muscle (skeletal); Vector: pDNR-LIB
(Clontech); Site_1: SfiI (ggccgctcgcc); Site_2: SfiI
(ggccattatggcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGGCGGCGCATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
BASE COUNT      256 a      183 c      181 g      223 t
ORIGIN

Query Match      100.0%; Score 81; DB 10; Length 843;
Best Local Similarity 100.0%; Pred. No. 2.6e-15;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 22:26:16 ; Search time 1960 Seconds
(without alignments)
1004.420 Million cell updates/sec

Title: US-09-585-077c-1_COPY_4300_4380
Perfect score: 81
Sequence: 1 gccacatcagactggctcaa.....aagaaggacagaatcccagc 81

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	699	12 BG761337	BG761337 602718213
2	81	100.0	706	10 BG616938	BG616938 602615612
3	81	100.0	843	10 BE971350	BE971350 601651514
4	80	98.8	490	9 AI065054	AI065054 HA0887 Hu

5	79.4	98.0	366	9	AV696724	AV696724
6	79.4	98.0	374	9	AV661012	AV661012
7	79.4	98.0	501	9	AI174790	AI174790 HA2511 Hu
8	79.4	98.0	540	12	BM821786	BM821786 K-EST0091
9	79.4	98.0	564	9	AV661502	AV661502
10	79.4	98.0	625	14	CB153827	CB153827 K-EST0211
11	79.4	98.0	636	9	AV654671	AV654671
12	79.4	98.0	637	9	AV683932	AV683932
13	79.4	98.0	646	9	AL046242	AL046242 DKF2p434E
14	79.4	98.0	949	10	BG533520	BG533520 601860935
15	78.4	96.8	362	9	AV656423	AV656423
16	77.8	96.0	365	9	AV660978	AV660978
17	77.8	96.0	626	9	AV653332	AV653332
18	77.8	96.0	830	10	BG569284	BG569284 602588576
19	74.6	92.1	444	9	AA883205	AA883205 am16h02.s
20	73.2	90.4	405	14	T59186	T59186 yb50e08.r1
21	73	90.1	350	14	CB781273	CB781273 AMGNNUC:C
22	73	90.1	554	12	BM432173	BM432173 IJEFJ1C10.
23	73	90.1	688	14	CB462893	CB462893 723232 MA
24	70.2	86.7	512	12	BQ011349	BQ011349 UI-1-BClp
25	69.8	86.2	473	9	AI786068	AI786068 uj57h08.Y
26	69.8	86.2	492	9	AA250015	AA250015 mz59d09.r
27	69.8	86.2	524	9	AA268939	AA268939 yb01903.r
28	69.8	86.2	629	9	AA986202	AA986202 uc73e03.Y
29	69.8	86.2	706	14	CB950475	CB950475 AGENCOURT
30	69.8	86.2	740	14	CB948817	CB948817 AGENCOURT
31	69.8	86.2	791	12	BI217440	BI217440 602933642
32	69.8	86.2	920	10	BF533080	BF533080 602073616
33	69.8	86.2	1057	14	W29382	W29382 mb99c06.r1
34	69.8	86.2	3161	11	AK028683	AK028683 Mus muscu
35	69	85.2	307	14	T77346	T77346 yd72e05.r1
36	69	85.2	521	9	AA190971	AA190971 zp83hl2.r
37	68.8	84.9	725	12	BI247566	BI247566 602960022
38	68.2	84.2	511	9	AA511517	AA511517 vj28c06.r
39	67.4	83.2	541	9	AA393420	AA393420 zt75f02.r
40	67.4	83.2	553	9	AA398737	AA398737 zt75f02.s
41	65.4	80.7	391	9	AV655328	AV655328
42	64	79.0	622	9	AI047399	AI047399 ud65g10.Y
43	60.4	74.6	564	14	CB120338	CB120338 K-EST0167
44	60	74.1	372	10	BF547163	BF547163 UI-R-C1-1
45	58.4	72.1	745	14	CB594152	CB594152 AGENCOURT

ALIGNMENTS

RESULT 1
BG761337
LOCUS BG761337 699 bp mRNA linear EST 15-MAY-2001
DEFINITION 602718213F1 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4858442 5',
mRNA sequence.
ACCESSION BG761337
VERSION BG761337.1 GI:14071990
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 699)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM1712 row: i column: 03
High quality sequence stop: 695.

QY 61 CAGAAGGACAGAATCCCAGC 81
| | | | | | | | | | | | | | | |
Db 96 CAGAAGGACAGAATCCCAGC 116

Search completed: September 12, 2003, 22:34:25
Job time : 211.5 secs

XX Toxicologically relevant human nucleotide sequence #1900.
DE
XX
KW Toxicologically relevant gene; toxicological response; gene; ss.
XX
OS Homo sapiens.
XX WO2003016500-A2.
XX
XX 27-FEB-2003.
XX
XX 16-AUG-2002; 2002WO-US26514.
XX
XX 16-AUG-2001; 2001US-313080P.
XX
XX (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
XX Neft RE, Dunn RT, Adkins K, Pickett GG, Kier LD, Schmeiser K;
PI Alen P;
XX
XX WPI; 2003-268322/26.
DR
XX
XX Determining a toxicological response to an agent, useful for screening
PT of drugs, comprises comparing the expression profile of one or more
PT human toxic response genes to a reference gene expression profile
PT indicative of toxicity -
XX
XX Claim 1; Page 447; 455pp; English.
XX
XX The present invention describes a method (M1) for determining a
CC toxicological response to an agent, which comprises comparing the
CC expression profile of one or more human toxic response genes to a
CC reference gene expression profile indicative of toxicity, and so
CC determining the presence of a toxic response to the agent. Also
CC described: (1) an array comprising one or more polynucleotides selected
CC from the genes corresponding to the partial sequences given in ABZ82842
CC to ABZ84764, or their fragments of at least 20 nucleotides, or
CC homologues; and (2) determining if a gene putatively identified to be a
CC toxic response gene plays a role on toxic response pathways by
CC determining the expression profile of the gene after exposure of cells
CC or a human subject to a known toxic pharmaceutical or industrial agent,
CC comprising: (a) exposing cells to an agent or isolating cells from a
CC human subject who was exposed to an agent; (b) obtaining the test gene
CC expression profile for a putatively identified toxic response gene after
CC exposure to a known toxic pharmaceutical or industrial agent; and
CC (c) comparing the test profile to the expression profile of a gene with
CC a similar function or comparing the test profile to the expression
CC profile of that gene after exposure to other known toxic compounds. The
CC methods are useful for predicting and determining toxicological responses
CC on a cellular, organ or system level. The arrays comprising the human
CC genes are useful for toxicological screening of drugs, pharmaceutical
CC compounds and chemicals.
XX
SQ Sequence 715 BP; 196 A; 173 C; 141 G; 205 T; 0 other;

Query Match 92.8%; Score 75.2; DB 25; Length 715;
Best Local Similarity 96.2%; Pred. No. 1e-16;
Matches 77; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACATCAGACTGGCTCAACGCCAACAAATGTCCTCGCAACCCAGTGGCATGGCGTCTC 61
|| |||||
Db 49 CCTATCAGACTGGCTCAACGCCAACAAATGTCCTCGCAACCCAGTGGCATGGCGTCTC 108

QY 62 AAGAGGACAGATCCCGC 81
|||||
Db 109 AAGAGGACAGATCCCGC 128

RESULT 15
ABX47313
ID ABX47313 standard; cDNA; 434 BP.
XX
AC ABX47313;

XX 21-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #12478.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX US2002137139-A1.
XX
XX 26-SEP-2002.
XX
XX 24-SEP-2001; 2001US-0960352.
XX
XX 12-JAN-1999; 99US-115707P.
PR 11-JAN-2000; 2000US-0480902.
XX
XX (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
XX WPI; 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and
PT analysis, cattle breeding, or for genetically improving cattle
XX
XX Claim 2; SEQ ID No 12478; 245pp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived
CC from cattle, and the LMFD nucleic acid can specifically hybridise to a
CC second nucleic acid molecule comprising any of 15112 nucleotide
CC sequences, appearing as ABX34836-ABX49947, or complements of them.
CC Also included are; (1) a transformed cell having a nucleic acid
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridisation between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMFD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome
CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMFD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139.
XX
SQ Sequence 434 BP; 135 A; 109 C; 86 G; 104 T; 0 other;

Query Match 90.1%; Score 73; DB 25; Length 434;
Best Local Similarity 93.8%; Pred. No. 5.4e-16;
Matches 76; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTCGCAACCCAGTGGCATGGCGTCT 60
|||||
Db 36 GCCACATCAGACTGGCTCAATGCCAACAAATGTCCTCGTACCCAGTGGCATGGCGTCT 95

Db 4370 CAAGAAGGACAGAAATCCACG 4390

RESULT 12

AAC89480

ID AAC89480 standard; DNA; 5761 BP.

XX AAC89480;

AC AAC89480;

XX 13-MAR-2001 (first entry)

DT Human T1405 carbamyl phosphate synthase I DNA.

DE Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;

XX sclerosis; pulmonary hypertension; bone marrow; gene therapy; ds.

KW Homo sapiens.

XX WO200073322-A1.

PN 07-DEC-2000.

PD 01-JUN-2000; 2000WO-US15079.

XX 01-JUN-1999; 99US-0323472.

PR (UYVA-) UNIV VANDERBILT.

XX Summar ML, Christman BW;

PI WPI; 2001-049926/06.

DR Detecting susceptibility to suboptimal urea cycle function, e.g. bone

XX marrow transplant toxicity, comprises identifying a polymorphism in the

PT gene for carbamylphosphate synthase -

PT Disclosure; Page 134-140; 171pp; English.

PS The present invention relates to screening for susceptibility to

XX suboptimal urea cycle function or to bone marrow transplant toxicity

CC by detecting a polymorphism in the carbamyl phosphate synthase I

CC (CPSI) gene. The method is used to detect subjects at risk of

CC hepatitis, sclerosis, pulmonary hypertension and bone marrow

CC transplant toxicity. These conditions can be treated or prevented

CC by administration of a nitric oxide precursor or by gene

CC therapy (administration of sequences that encode CPSI).

XX Sequence 5761 BP; 1624 A; 1210 C; 1309 G; 1618 T; 0 other;

SQ Query Match 98.0%; Score 79.4; DB 22; Length 5761;

Best Local Similarity 98.8%; Pred. No. 5.4e-18;

Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCCTGCCAACCCAGTGGCATGGCGTCT 60

Db 4300 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCCTGCCAACCCAGTGGCATGGCGTCT 4359

QY 61 CAAGAAGGACAGAAATCCACG 81

Db 4360 CAAGAAGGACAGAAATCCACG 4380

RESULT 13

AB235442

ID AB235442 standard; cDNA; 5772 BP.

XX AB235442;

AC AB235442;

XX 05-FEB-2003 (first entry)

DT Human gene expression profile polynucleotide SEQ ID NO 553.

DE Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;

XX Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;

KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;

KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;

KW gene expression; gene; ss.

XX Homo sapiens.

OS WO200274979-A2.

PN 26-SEP-2002.

XX 20-MAR-2002; 2002WO-US08456.

PF 20-MAR-2001; 2001US-276947P.

PR (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.

XX Wan J, Wang Y;

PI WPI; 2002-740862/80.

DR New gene expression profile generated from primary, endothelial,

XX epithelial, and muscle cell types, useful for identifying disease

PT pathologies involving alterations of gene expression, e.g. cancer

PT Disclosure; Page 691-693; 850pp; English.

XX The invention relates to a gene expression profile comprising one or more

CC genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type

CC is a coronary artery endothelium, umbilical artery or vein endothelium,

CC aortic endothelium, dermal microvascular endothelium, pulmonary artery

CC endothelium, myometrium microvascular endothelium, keratinocyte

CC epithelium, bronchial epithelium, mammary epithelium, prostate

CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,

CC small airway epithelium, renal epithelium, umbilical artery smooth

CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,

CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,

CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,

CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,

CC osteoblasts or prostate stromal cell. The gene expression profile is used

CC for determining the level of RNA expression for a sample, determining the

CC phenotype of a cell and distinguishing cell types. The gene or a protein

CC expression profile is useful in identifying disease pathologies

CC involving alterations of gene expression. The assessment of expression

CC profiles may provide meaningful information with respect to tumour type

CC and stage, treatment methods, and prognosis. The gene or protein

CC expression profile may also be used for creating microarrays. The

CC microarray is useful for genetic and physical mapping of genomes, DNA

CC sequencing, genetic or medical diagnosis, genotyping of organisms,

CC confirming cell or tissue identifications and in identifying promising

CC antibiotics, antiviral or antifungal agents.

XX Sequence 5772 BP; 1602 A; 1180 C; 1295 G; 1571 T; 124 other;

SQ Query Match 98.0%; Score 79.4; DB 24; Length 5772;

Best Local Similarity 98.8%; Pred. No. 5.4e-18;

Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCCTGCCAACCCAGTGGCGTCT 60

Db 4310 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCCTGCCAACCCAGTGGCGTCT 4369

QY 61 CAAGAAGGACAGAAATCCACG 81

Db 4370 CAAGAAGGACAGAAATCCACG 4390

RESULT 14

ABZ84741

ID ABZ84741 standard; cDNA; 715 BP.

XX ABZ84741;

AC ABZ84741;

XX 14-MAY-2003 (first entry)

Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCAGTGGCATGGCCGTCT 216

QY 61 CAAGAAGGACAGAAATCCAGC 81
|||||

Db 217 CAAGAAGGACAGAAATCCAGC 237

RESULT 10
ABX99366
ID ABX99366 standard; cDNA; 531 BP.
XX
AC ABX99366;
XX
DT 22-MAY-2003 (first entry)
XX
DE Lung cancer therapyand diagnosis associated cDNA #93.
XX
KW Lung cancer; cytostatic; vaccine; gene therapy; cancer;
KW gene; ss.
KW
OS Homo sapiens.
XX
PN US2002172952-A1.
XX
PD 21-NOV-2002.
XX
PF 10-JUL-2001; 2001US-0902941.
XX
PR 30-JUN-1999; 99US-0346492.
PR 15-OCT-1999; 99US-0419356.
PR 17-DEC-1999; 99US-0466867.
PR 30-DEC-1999; 99US-0476300.
PR 06-MAR-2000; 2000US-0519642.
PR 22-MAR-2000; 2000US-0533077.
PR 10-APR-2000; 2000US-0546259.
PR 27-APR-2000; 2000US-0560406.
PR 05-JUN-2000; 2000US-0589184.
PR 11-JUL-2000; 2000US-0614124.
PR 29-AUG-2000; 2000US-0651563.
PR 08-SEP-2000; 2000US-0658824.
PR 26-SEP-2000; 2000US-0671325.
PR 06-OCT-2000; 2000US-0677419.
PR 30-OCT-2000; 2000US-0702705.
PR 13-DEC-2000; 2000US-0736457.
PR 03-MAY-2001; 2001US-0849626.
XX
PA (CORI-) CORIXA CORP.
XX
PI Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW, Durham M;
PI Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;
XX
DR WPI; 2003-328427/31.
XX
PT New polynucleotide, useful for preparing a composition for treating or
PT inhibiting development of cancer, e.g. lung cancer -
XX
PS Example 1; SEQ ID NO 93; 82pp; English.
XX
CC The invention describes an isolated polynucleotide comprising one of 32
CC sequences, complement or degenerate variants of them. The polynucleotide
CC is useful for preparing a composition e.g. a vaccine or for gene therapy,
CC for treating or inhibiting development of cancer, e.g. lung cancer.
CC This sequence represents a polynucleotide associated with the
CC compositions and methods for the therapy and diagnosis of lung cancer.
XX
SQ Sequence 531 BP; 157 A; 139 C; 108 G; 127 T; 0 other;

Query Match 98.0%; Score 79.4; DB 25; Length 531;
Best Local Similarity 98.8%; Pred. No. 3.2e-18;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCAGTGGCATGGCCGTCT 216

QY 61 CAAGAAGGACAGAAATCCAGC 81
|||||

Db 217 CAAGAAGGACAGAAATCCAGC 237

RESULT 11
AAH57467
ID AAH57467 standard; cDNA; 5195 BP.
XX
AC AAH57467;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human liver cell specific cDNA sequence SEQ ID NO:307.
XX
KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
OS Homo sapiens.
XX
PN WO200132927-A2.
XX
PD 10-MAY-2001.
XX
PF 02-NOV-2000; 2000WO-US30396.
XX
PR 04-NOV-1999; 99US-0163508.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Sornasse T, Seilhamer JJ, Watson GA;
XX
DR WPI; 2001-291057/30.
XX
PT New cell and tissue specific polynucleotides useful for diagnosis,
PT prognosis or monitoring of treatments for disorders where the gene is
PT associated with a cancer, immunopathology or neuropathology -
XX
PS Claim 1; Page 230-231; 327pp; English.
XX
CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC sequences (I). (I) can have cytostatic, immunomodulatory and
CC neuroprotective activities, and can be used in gene therapy. (I) and
CC proteins (II) encoded by then are used in high throughput screening
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC their fragments, immunoglobulins, inhibitors, drug compounds and
CC pharmaceutical agents. Expression of (I) in a sample indicates the
CC differentiation of embryonic stem cells into a tissue selected from
CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC tissues. (I) and (II) are used to produce an expression profile that
CC defines a metabolic or developmental process, treatment, condition,
CC disease or disorder. The gene profile can be used for diagnosis,
CC prognosis or monitoring of treatments and for investigating a
CC predisposition to a disorder where the gene is associated with a
CC cancer, immunopathology or neuropathology.
XX
SQ Sequence 5195 BP; 1469 A; 1102 C; 1211 G; 1413 T; 0 other;

Query Match 98.0%; Score 79.4; DB 22; Length 5195;
Best Local Similarity 98.8%; Pred. No. 5.3e-18;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCAGTGGCATGGCCGTCT 60
|||||

Db 4310 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCACCCAGTGGCATGGCCGTCT 4369

QY 61 CAAGAAGGACAGAAATCCAGC 81
|||||

PF 10-JUL-2001; 2001WO-US22058.
XX
PR 11-JUL-2000; 2000US-0614124.
PR 29-AUG-2000; 2000US-0651563.
PR 08-SEP-2000; 2000US-0658824.
PR 26-SEP-2000; 2000US-0671325.
PR 06-OCT-2000; 2000US-0677419.
PR 30-OCT-2000; 2000US-0702705.
PR 13-DEC-2000; 2000US-0736457.
PR 03-MAY-2001; 2001US-0849626.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;
PI McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;
XX
DR WPI; 2002-164634/21.
XX
PT Novel polynucleotide encoding a lung tumour polypeptide useful for
PT stimulating and/or expanding T cells specific for a tumour protein -
XX
PS Example 1; SEQ ID No 93; 223pp; English.
XX
CC The invention describes an isolated polynucleotide and polypeptide
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein for determining the presence of a cancer in a patient. A
CC composition containing the polynucleotide and/or polypeptide is useful
CC for treating a lung cancer in a patient. The polypeptide is useful for
CC removing tumour cells from a biological sample. The polynucleotide is
CC also useful as probe or primer to detect the level of mRNA encoding a
CC tumour protein. This sequence encodes a lung tumour associated protein
CC or protein fragment, described in the method of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 531 BP; 157 A; 139 C; 108 G; 127 T; 0 other;

Query Match 98.0%; Score 79.4; DB 24; Length 531;
Best Local Similarity 98.8%; Pred. No. 3.2e-18;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGCGCTCT 60
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGCGCTCT 216

QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 217 CAAGAAGGACAGAATCCCAGC 237

RESULT 9
ACAL0415
ID ACAL0415 standard; cDNA; 531 BP.
XX
AC ACAL0415;
XX
DT 05-JUN-2003 (first entry)
XX
DE Human lung cancer-associated cDNA, SEQ ID 93.
DE
KW Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW T cell expansion; CD4; CD8.
XX
OS Homo sapiens.
XX
PN US2002197669-A1.
XX
PD 26-DEC-2002.
XX
PF 03-MAY-2001; 2001US-0849626.

XX 13-DEC-2000; 2000US-0736457.
PR
PA (BANG/) BANGUR C S.
PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.
XX
PI Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;
PI Clapper JD;
XX
DR WPI; 2003-352750/33.
XX
PT Novel lung cancer polynucleotide encoding lung cancer protein, useful
PT for detecting the presence of lung cancer in a patient, and in
PT pharmaceutical compositions, e.g. vaccines, for treating lung cancer -
XX
PS Example 1; Page -: 72pp; English.
XX
CC The invention relates to a polynucleotide encoding a lung tumour protein,
CC comprising a sequence selected from any of the 14 sequences
CC mentioned in the specification, or a sequence (S2) mentioned in
CC specification, complement of S1, sequences consisting of at least 20
CC contiguous residues of S1, sequences that hybridise to S1, sequences
CC having 75%, preferably 90%, identity to S1, or degenerate variants of
CC S1. Also included are an isolated polypeptide (comprising a sequence (S3)
CC selected from any one of the 4 amino acid sequences mentioned in the
CC specification, a sequence encoded by the polynucleotide, or sequences
CC having at least 70%, preferably 90%, identity to a sequence encoded by
CC the polynucleotide), an expression vector comprising the polynucleotide
CC operably linked to an expression control sequence, a host cell
CC transformed or transfected with the vector, an isolated antibody (or its
CC antigen-binding fragment) that specifically binds to the polypeptide,
CC detecting the presence of a cancer in a patient, a fusion protein
CC comprising the polypeptide, an oligonucleotide that hybridises to
CC S1 under moderately stringent conditions, stimulating and/or expanding T
CC cells specific for a tumour protein (comprising contacting T cells with
CC the polynucleotide, protein or antigen-presenting cells, under conditions
CC and for a time sufficient to permit the stimulation and/or expansion of T
CC cells) and inhibiting the development of a cancer in a patient (by
CC incubating CD4+ and/or CD8+ T cells isolated from a patient with the
CC polynucleotide, protein or antigen presenting cells that express the
CC polynucleotide, such that T cells proliferate, administering to the
CC patient an effective amount of the proliferated T cells, and thus
CC inhibiting the development of a cancer in the patient. The
CC polynucleotide, protein and cells are useful in a composition for
CC stimulating an immune response in a patient, and for treating a cancer in
CC a patient (particularly lung cancer). The oligonucleotide is useful for
CC determining the presence of a cancer in a patient. The protein and
CC oligonucleotides are useful in pharmaceutical compositions, e.g.
CC vaccines. The polynucleotide is also useful as a probe or primer for
CC nucleic acid hybridisation, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. An amplified portion of the polynucleotide is
CC useful for isolating a full-length gene from a suitable library.
CC The present sequence is a cDNA (full length, extended or partial)
CC isolated from a library derived from lung tumour/cancer cells.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from the USPTO
CC at seqdata.uspto.gov/sequence.html?DocId=20020197669.
XX
SQ Sequence 531 BP; 157 A; 139 C; 108 G; 127 T; 0 other;

Query Match 98.0%; Score 79.4; DB 25; Length 531;
Best Local Similarity 98.8%; Pred. No. 3.2e-18;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGCGCTCT 60
|||||
GCGCGCTCT 60
|||||

PN WO200073322-A1.
XX 07-DEC-2000.
PD
XX 01-JUN-2000; 2000WO-US15079.
PF
XX 01-JUN-1999; 99US-0323472.
PR
XX (UYVA-) UNIV VANDERBILT.
PA
XX Summar ML, Christman BW;
PI
XX WPI; 2001-049926/06.
DR
XX Detecting susceptibility to suboptimal urea cycle function, e.g. bone
PT marrow transplant toxicity, comprises identifying a polymorphism in the
PT gene for carbamylphosphate synthase -
XX
PS Disclosure; Fig 10; 171pp; English.
XX
CC The present invention relates to screening for susceptibility to
CC suboptimal urea cycle function or to bone marrow transplant toxicity
CC by detecting a polymorphism in the carbamyl phosphate synthase I
CC (CPSI) gene. The method is used to detect subjects at risk of
CC hepatitis, sclerosis, pulmonary hypertension and bone marrow
CC transplant toxicity. These conditions can be treated or prevented
CC by administration of a nitric oxide precursor or by gene
CC therapy (administration of sequences that encode CPSI).
XX
SQ Sequence 495 BP; 158 A; 96 C; 85 G; 155 T; 1 other;

Query Match 98.0%; Score 79.4; DB 22; Length 495;
Best Local Similarity 98.8%; Pred. No. 3.1e-18;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGTGCCAACCCAGTGGCGGTCT 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
126 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGTGCCAACCCAGTGGCGGTCT 185

QY 61 CAAGAAGGACAGAAATCCGAGC 81
Db ||||||||||||||||||||||||||||
186 CAAGAAGGACAGAAATCCGAGC 206

RESULT 7
AAF68175
ID AAF68175 standard; cDNA; 531 BP.
XX
AC AAF68175;
XX
DT 12-APR-2001 (first entry)
XX
DE Human lung tumour protein related nucleotide sequence SEQ ID NO:93.
XX
KW Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
KW cytostatic; antisense inhibition; ss.
XX
OS Homo sapiens.
XX
PN WO200100828-A2.
XX
PD 04-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18061.
XX
PR 30-JUN-1999; 99US-0346492.
PR 15-OCT-1999; 99US-0419356.
PR 17-DEC-1999; 99US-0466867.
PR 30-DEC-1999; 99US-0476300.
PR 06-MAR-2000; 2000US-0519642.
PR 22-MAR-2000; 2000US-0533077.
PR 10-APR-2000; 2000US-0546259.

PR 27-APR-2000; 2000US-0560406.
PR 05-JUN-2000; 2000US-0589184.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
PI Retter MW, Mannion J;
XX
DR WPI; 2001-071488/08.
XX
PT Lung tumor-associated proteins and the nucleic acids that encode them,
PT useful for preventing, diagnosing and treating lung cancer -
XX
PS Example 1; Page 181; 436pp; English.
XX
CC The present invention describes immunogenic portions of lung tumour-
CC associated proteins (I) and the nucleic acids (NAs) that encode them.
CC (I) have cytostatic activity and can be used in gene therapy, antisense
CC inhibition and in vaccines. The NAs and the lung tumour-associated
CC proteins they encode may be used in the prevention, treatment and
CC diagnosis of diseases associated with their inappropriate expression,
CC especially lung cancers. For example, the NAs may be administered to
CC treat diseases by rectifying mutations or deletions in a patient's genome
CC that affect the activity of the protein by expressing inactive proteins
CC or to supplement the patients own production of (I). Additionally, the
CC NAs may be used to produce the lung-tumour associated protein, according
CC to standard recombinant DNA methodology. Conversely, antisense NA
CC molecules may be administered to down regulate protein expression by
CC binding with the cells own genes and preventing their expression. The NA
CC and complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar NA sequences in
CC samples, and hence which patients may be in need of treatment for lung
CC cancer. The (I) may be used as antigens in the production of antibodies
CC and in assays to identify modulators (agonists and antagonists) of the
CC expression and activity of the protein. AAF68083 to AAF68878 and
CC AAB76848 to AAB76878 represent human lung tumour protein related
CC nucleotide and protein sequences which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 531 BP; 157 A; 139 C; 108 G; 127 T; 0 other;

Query Match 98.0%; Score 79.4; DB 22; Length 531;
Best Local Similarity 98.8%; Pred. No. 3.2e-18;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGTGCCAACCCAGTGGCGGTCT 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGTGCCAACCCAGTGGCGGTCT 216

QY 61 CAAGAAGGACAGAAATCCGAGC 81
Db ||||||||||||||||||||||||
217 CAAGAAGGACAGAAATCCGAGC 237

RESULT 8
ABK38086
ID ABK38086 standard; cDNA; 531 BP.
XX
AC ABK38086;
XX
DT 21-MAY-2002 (first entry)
XX
DE cDNA encoding clone #19065 of lung tumour protein.
XX
KW Lung tumour; cancer; T cell; immune response stimulator;
KW cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200204514-A2.
XX
PD 17-JAN-2002.
XX

Query Match 100.0%; Score 81; DB 22; Length 5761;
Best Local Similarity 100.0%; Pred. No. 1.5e-18;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCCTGCGCAACCCAGTGGCGGTCT 60
|||||
Db 4300 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCCTGCGCAACCCAGTGGCGGTCT 4359

QY 61 CAAGAAGGACAGAATCCCAGC 81
|||||
Db 4360 CAAGAAGGACAGAATCCCAGC 4380

RESULT 4
AAC89487
ID AAC89487 standard; DNA; 5761 BP.
XX
AC AAC89487;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human carbamyl phosphate synthase DNA.
XX
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200073322-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15079.
XX
PR 01-JUN-1999; 99US-0323472.
XX
PA (UYVA-) UNIV VANDERBILT.
PI Summar ML, Christman BW;
XX
XX WPI; 2001-049926/06.
XX
PT Detecting susceptibility to suboptimal urea cycle function, e.g. bone
PT marrow transplant toxicity, comprises identifying a polymorphism in the
PT gene for carbamylphosphate synthase -
XX
PS Claim 65; Page 146; 171pp; English.
XX
CC The present invention relates to screening for susceptibility to
CC suboptimal urea cycle function or to bone marrow transplant toxicity
CC by detecting a polymorphism in the carbamyl phosphate synthase I
CC (CPSI) gene. The method is used to detect subjects at risk of
CC hepatitis, sclerosis, pulmonary hypertension and bone marrow
CC transplant toxicity. These conditions can be treated or prevented
CC by administration of a nitric oxide precursor or by gene
CC therapy (administration of sequences that encode CPSI).
XX
SQ Sequence 5761 BP; 1625 A; 1209 C; 1309 G; 1618 T; 0 other;

Query Match 100.0%; Score 81; DB 22; Length 5761;
Best Local Similarity 100.0%; Pred. No. 1.5e-18;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCCTGCGCAACCCAGTGGCGGTCT 60
|||||
Db 4300 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCCTGCGCAACCCAGTGGCGGTCT 4359

QY 61 CAAGAAGGACAGAATCCCAGC 81
|||||
Db 4360 CAAGAAGGACAGAATCCCAGC 4380

RESULT 5

AAC89488
ID AAC89488 standard; DNA; 5761 BP.
XX
AC AAC89488;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human carbamyl phosphate synthase DNA.
XX
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200073322-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15079.
XX
PR 01-JUN-1999; 99US-0323472.
XX
PA (UYVA-) UNIV VANDERBILT.
PI Summar ML, Christman BW;
XX
XX WPI; 2001-049926/06.
XX
PT Detecting susceptibility to suboptimal urea cycle function, e.g. bone
PT marrow transplant toxicity, comprises identifying a polymorphism in the
PT gene for carbamylphosphate synthase -
XX
PS Claim 65; Page 156-161; 171pp; English.
XX
CC The present invention relates to screening for susceptibility to
CC suboptimal urea cycle function or to bone marrow transplant toxicity
CC by detecting a polymorphism in the carbamyl phosphate synthase I
CC (CPSI) gene. The method is used to detect subjects at risk of
CC hepatitis, sclerosis, pulmonary hypertension and bone marrow
CC transplant toxicity. These conditions can be treated or prevented
CC by administration of a nitric oxide precursor or by gene
CC therapy (administration of sequences that encode CPSI).
XX
SQ Sequence 5761 BP; 1626 A; 1209 C; 1308 G; 1618 T; 0 other;

Query Match 100.0%; Score 81; DB 22; Length 5761;
Best Local Similarity 100.0%; Pred. No. 1.5e-18;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCCTGCGCAACCCAGTGGCGGTCT 60
|||||
Db 4300 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCCTGCGCAACCCAGTGGCGGTCT 4359

QY 61 CAAGAAGGACAGAATCCCAGC 81
|||||
Db 4360 CAAGAAGGACAGAATCCCAGC 4380

RESULT 6
AAC89481
ID AAC89481 standard; DNA; 495 BP.
XX
AC AAC89481;
XX
DT 13-MAR-2001 (first entry)
XX
DE Target 5' and 3' sequence.
XX
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy; ds.
XX
OS Homo sapiens.
XX

DR WPI; 1993-006237/01.
DR P-PSDB; AAR30636.
XX
PT Carbamyl phosphate synthetase I gene - used to detect diseases
PT caused by carbamyl phosphatase synthetase by using overlapped
PT cDNA contg. total human cDNA sequence
XX
PS Claim 1; Page 6-8; 12pp; Japanese.
XX
CC This sequence is the human carbamyl phosphate synthetase I (CPSI)
CC gene. This sequence or fragments of it may be used as probes to
CC detect mutations in the CPSI gene. CPSI deficiency disease may be
CC detected by using overlapping cDNA representing the full length cDNA
CC sequence of human CPSI. This sequence was isolated using three
CC amplified fragments derived from protein coding regions of the rat
CC CPSI cDNA as probes.
XX
SQ Sequence 5215 BP; 1484 A; 1104 C; 1205 G; 1422 T; 0 other;

Query Match 100.0%; Score 81; DB 14; Length 5215;
Best Local Similarity 100.0%; Pred. No. 1.4e-18;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCAACCCAGTGGCAGCGCGTCT 60
|||||
Db 4295 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCAACCCAGTGGCAGCGCGTCT 4354

QY 61 CAAGAAGGACAGAAATCCGAGC 81
|||||
Db 4355 CAAGAAGGACAGAAATCCGAGC 4375

RESULT 2
ABN95271
ID ABN95271 standard; DNA; 5215 BP.
XX
AC ABN95271;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #1769 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
PS Claim 1; SEQ ID NO 1769; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of

CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 5215 BP; 1484 A; 1102 C; 1207 G; 1422 T; 0 other;

Query Match 100.0%; Score 81; DB 24; Length 5215;
Best Local Similarity 100.0%; Pred. No. 1.4e-18;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCAACCCAGTGGCAGCGCGTCT 60
|||||
Db 4295 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCCTGCCAACCCAGTGGCAGCGCGTCT 4354

QY 61 CAAGAAGGACAGAAATCCGAGC 81
|||||
Db 4355 CAAGAAGGACAGAAATCCGAGC 4375

RESULT 3
AAC89479
ID AAC89479 standard; DNA; 5761 BP.
XX
AC AAC89479;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human N1405 carbamyl phosphate synthase I DNA.
XX
KW Urea cycle; carbamyl phosphate synthase I; CPSI; hepatitis;
KW sclerosis; pulmonary hypertension; bone marrow; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200073322-A1.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15079.
XX
PR 01-JUN-1999; 99US-0323472.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Summar ML, Christman BW;
XX
DR WPI; 2001-049926/06.
XX
XX
PT Detecting susceptibility to suboptimal urea cycle function, e.g. bone
PT marrow transplant toxicity, comprises identifying a polymorphism in the
PT gene for carbamylphosphate synthase -
XX
PS Disclosure; Page 124-130; 171pp; English.
XX
XX
CC The present invention relates to screening for susceptibility to
CC suboptimal urea cycle function or to bone marrow transplant toxicity
CC by detecting a polymorphism in the carbamyl phosphate synthase I
CC (CPSI) gene. The method is used to detect subjects at risk of
CC hepatitis, sclerosis, pulmonary hypertension and bone marrow
CC transplant toxicity. These conditions can be treated or prevented
CC by administration of a nitric oxide precursor or by gene
CC therapy (administration of sequences that encode CPSI).
XX
SQ Sequence 5761 BP; 1625 A; 1209 C; 1309 G; 1618 T; 0 other;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 22:19:36 ; Search time 201.5 Seconds
(without alignments)
1085.134 Million cell updates/sec

Title: US-09-585-077C-1_COPY_4300_4380
Perfect score: 81
Sequence: 1 gccacatcagactggctcaa.....aagaaggacagaatccaccg 81

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	5215	14	AAQ34768
2	81	100.0	5215	24	ABN95271
3	81	100.0	5761	22	AAC89479
4	81	100.0	5761	22	AAC89487
5	81	100.0	5761	22	AAC89488
6	79.4	98.0	495	22	AAC89481
7	79.4	98.0	531	22	AAF68175
8	79.4	98.0	531	24	ABK38086
					hCPSI gene. Homo
					Gene #1769 used to
					Human N1405 carbam
					Human carbamyl pho
					Human carbamyl pho
					Target 5' and 3' s
					Human lung tumour
					cDNA encoding clon

9	79.4	98.0	531	25	ACA10415	Human lung cancer-
10	79.4	98.0	531	25	ABX99366	Lung cancer therap
11	79.4	98.0	5195	22	AAH57467	Human liver cell s
12	79.4	98.0	5761	22	AAC89480	Human T1405 carbam
13	79.4	98.0	5772	24	ABZ35442	Human gene express
14	75.2	92.8	715	25	ABZ84741	Toxicologically re
15	73	90.1	434	25	ABX47313	Bovine EST associa
16	69	85.2	816	24	ABA93013	Human cDNA clone B
17	62.8	77.5	459	22	AAH57259	Human liver specif
18	46	56.8	580	24	ABT08945	Phase-1 Rat CT gen
19	40.2	49.6	242	25	ABX36653	Bovine EST associa
20	31	38.3	31	22	AAI31280	Human single nucle
21	29.4	36.3	3591	22	AAI02819	Human reproductive
22	27.4	33.8	646	24	ABT09297	Phase-1 Rat CT gen
23	26.8	33.1	8098	22	AAK90713	Human digestive sy
24	26.8	33.1	9556	22	AAK90714	Human digestive sy
25	26.6	32.8	2372	8	AAH70814	Aspergillus niger
26	26.4	32.6	2153	24	ABI99890	Mouse ischaemic co
27	26.4	32.6	2168	18	AAV00499	Murine alpha-(2) s
28	26.2	32.3	6564	22	AAF31637	Mycobacterium tube
29	26.2	32.3	4403765	22	AAI99683	Mycobacterium tube
30	26.2	32.3	4411529	22	AAI99682	Mycobacterium tube
31	26	32.1	375	22	AAH36700	Mycobacterium tube
32	26	32.1	439	22	AAI88303	Human colon cancer
33	26	32.1	987	21	AAC79247	Human polynucleoti
34	26	32.1	987	23	AAD23323	Human lung tumour-
35	26	32.1	1987	20	AAZ53006	Human lung tumour-
36	26	32.1	2236	20	AAZ41294	Human prostate tum
37	25.8	31.9	349980	22	AAH41225	Human normal ovari
38	25.8	31.9	349980	22	AAH41226	Pyrococcus abyssi
39	25.6	31.6	799	22	AAI93930	Pyrococcus abyssi
40	25.6	31.6	1036	24	ABK48308	Human neuroblastom
41	25.6	31.6	1036	25	ACA56710	Human inhibitory G
42	25.6	31.6	1202	19	AAV71981	Human signalling p
43	25.4	31.4	288	25	ABX54929	A nucleotide seque
44	25.4	31.4	26225	22	AAAS35776	Bovine EST associa
45	25	30.9	25	25	ABZ84401	Human cardiovascular
						Toxicologically re

ALIGNMENTS

RESULT 1	
AAQ34768	
ID	AAQ34768 standard; DNA; 5215 BP.
XX	
AC	AAQ34768;
XX	
DT	06-MAY-1993 (first entry)
XX	
DE	hCPSI gene.
XX	
KW	Human; carbamyl phosphate synthetase I; CPSI; probe; mutation;
KW	CPSI deficiency disease; rat; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
XX	119..4621
XX	/*tag= a
PN	JP04335889-A.
XX	
PD	24-NOV-1992.
XX	
PF	09-MAY-1991; 91JP-0135902.
XX	
PR	09-MAY-1991; 91JP-0135902.
XX	
PA	(HARA/) HARAGUCHI Y.
PA	(MATS/) MATSUDA I.
XX	(MORI/) MORI M.
XX	

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 157 a 139 c 108 g 127 t
ORIGIN

Query Match 98.0%; Score 79.4; DB 6; Length 531;
Best Local Similarity 98.8%; Pred. No. 4.7e-17;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
|||||
Db 157 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 216
|||||

QY 61 CAAGAAGGACAGAAATCCCAGC 81
|||||
Db 217 CAAGAAGGACAGAAATCCCAGC 237
|||||

RESULT 12
AX367383
LOCUS AX367383 531 bp DNA linear PAT 17-FEB-2002
DEFINITION Sequence 93 from Patent WO2024514.
ACCESSION AX367383
VERSION AX367383.1 GI:18855487
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W.,
Marnerakis,M., Carter,D., Fanger,G.R., Vedvick,T.S., Bangur,C.S.,
McNabb,A., Fanger,N., Switzer,A., Mcneill,P.D. and Clapper,J.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0204514-A 93 17-JAN-2002;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..531
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 157 a 139 c 108 g 127 t
ORIGIN

Query Match 98.0%; Score 79.4; DB 6; Length 531;
Best Local Similarity 98.8%; Pred. No. 4.7e-17;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
|||||
Db 157 GCCACATCAGACTGGCTCAACGCCAACAAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 216
|||||

QY 61 CAAGAAGGACAGAAATCCCAGC 81
|||||
Db 217 CAAGAAGGACAGAAATCCCAGC 237
|||||

RESULT 13
HSCPS1MUT
LOCUS HSCPS1MUT 4503 bp mRNA linear PRI 24-SEP-1998
DEFINITION Homo sapiens mRNA for carbamyl phosphate synthetase I.
ACCESSION Y15793
VERSION Y15793.1 GI:3228247
KEYWORDS carbamoyl-phosphate synthetase I; CPS1 gene; mutation.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Finckh,U., Kohlschutter,A., Schafer,H., Sperhake,K., Colombo,J.P.
and Gal,A.

TITLE Prenatal diagnosis of carbamoyl phosphate synthetase I deficiency
by identification of a missense mutation in CPS1
JOURNAL Hum. Mutat. 12 (3), 206-211 (1998)
MEDLINE 98375696
PUBMED 9711878
REFERENCE 2
AUTHORS Finckh,U.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-1997) U. Finckh, Dept. of Human Genetics,
University Hospital Eppendorf, Butenfeld 42, 22529 Hamburg, FRG
REMARK revised by [3]
REFERENCE 3 (bases 1 to 4503)
AUTHORS Finckh,U.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1998) U. Finckh, Dept. of Human Genetics,
University Hospital Eppendorf, Butenfeld 42, 22529 Hamburg, FRG
COMMENT On Jun 16, 1998 this sequence version replaced gi:2959899.
Related sequence D90282.
FEATURES
source Location/Qualifiers
1..4503
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="q35"
/tissue_type="liver"
1..4503
/gene="CPS1"
1..4503
/gene="CPS1"
/codon_start=1
/product="carbamoyl phosphate synthetase 1"
/protein_id="CAA75785.1"
/db_xref="GI:3228248"
/db_xref="SWISS-PROT:P31327"
/translation="MTRLFAFKVVRTLKTGFGTNTVTAHQKWKFSRPGIRLLSVKRAQ
TAHIVLEDGCTMKGYSGFHESSVAGEVVFNTGLGGYPEAITDPAYKGQILTMANPIIG
NGGAPDITALDELGLSKYLESGIKVSGLLVLDYSKDYNHWLATKSLGOWLQEEKVPA
IYGDTRMLTKIIRDKGTMLCKIEFEGQPVDFDPNKQNLIAEVSTKDVKYGKGNPT
KVAVDCGIGKNNVIRLLVKGAEVHLVPWNHDTTKMEYDGLIAGGPGPALAEPLIQ
NVRKILESDRKEPLFGISTGNLITGLAAGAKTYKMSMANRGQNQPVNLITNKQAFITA
QNHGYALDNTLPAGWKPLFVNVDQTNNEGIMHESKPFPAVGFHPEVTPGIDTEYLFQ
SFFSLIKKGKATITISVLPKPAIVASRVEVSKVLILGSGCLSIGQAGEFDYSGSQAVK
AMKEENVKTVLMNPNIASVQINEVGLKQADTVYFLPITPQFVTEVIKAFQPDGLILGM
GGQALNCGVELFKRGVLKEYGVKVLGTSVESIMATEDRQLFSDKLAINEKIAFPSA
VESIEDALKAADTIGYPVMIRSAYALGGLSGICPNREILMDLSTKAFAMTNQILVEK
SVTGWKEIEYEVVRDADDNCTVTCNMENVNDAMGVHTGDSVVVAPAQILSNAEFQMLRR
TSINVVRLGIVGECNIQFALHPTSMEYCIIEVNARLSRSSALASKATGYPLAFIAAK
IALGIPLPEIKNVVSGKTSACFEPSLDYMTKIPRWDLDRFHGTSSRIGSSMKSVGEV
MAIGRTFEESQKALRMCHPSIEGTPRLPMNKWEPNSLDLRKELSEPSSTRIYAIK
AIDNMSLDEIEKLTYIDKWFLYKMRDILNMEKTLKGLNSESMTETLKRKEIGFSD
KQISKCLGLTEAQTRRLKKNHPWKQIDTLAAEYPSVTNLYVTYNGQEHVNFQ
DHGMVILGCGPVHIGSSVEFDWCAVSSIRTLRQLGKKTVVNCNPETVSTDFDECCKL
YFEELSLERILDYHQEACGGCITSVGGQIPNNLAVPLYKNGVKIMGTSPLQIDRAED
RSIFSAVLDELKVAQAPWKAVENTLNEALEFAKSDVDPCLLRPSYVLSCSAMNVVFSQ
EMKKFLEEATRVSQEHPVLTKEVEGAREVEMDAVGKGRVISHAISEHVEDAGVHSG
DATLMLPTQTISQGAIEKVKDATRKIAKAPAISSGPNVQFLVKGNDVILVIECNLRASR
SFPFVSKTLGVDFIDVATKVMIGENVDEKHLPTLDHPITPADYVAIKAPMFSWPRLRD
ADPILRCEMASTGEVACFGEIGHTAFLKAMLSTGFKIFQKILIGIQOFSRFRFLGVA
EQLHNEGFKLFATEATSDWLNANNVPATPVAMPVPSQEQGNPSLSSIRKLIRDSIDLVI
NLPNNNTKVFHDNVVIRRTAVDSGIPLLTNFQVTKLFAEAVQKSRKVDKSLFHYROY
SAGKAA"
1631
/gene="CPS1"
/note="mutation in CPS1 deficiency"
/replace="t"
2679
/gene="CPS1"
/note="putative polymorphism (silent)"
2838
/gene="CPS1"
/note="ambiguous"

variation

variation

variation

```

/db_xref="taxon:9606"
/chromosome="2"
/map="2q35"
84..196
/gene="CPSI"
/number=36
93 a 64 c 48 g 91 t
BASE COUNT
ORIGIN
exon
Query Match 98.0%; Score 79.4; DB 9; Length 296;
Best Local Similarity 98.8%; Pred. No. 5e-17;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
Db 99 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 158
QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 159 CAAGAAGGACAGAATCCCAGC 179
RESULT 8
AR184418
LOCUS AR184418 495 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 5 from patent US 6346382.
ACCESSION AR184418
VERSION AR184418.1 GI:20230383
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 495)
AUTHORS Summar,M.L. and Christman,B.W.
TITLE Human carbamyl phosphate synthetase I polymorphism and diagnostic methods related thereto
JOURNAL Patent: US 6346382-A 5 12-FEB-2002;
FEATURES Location/Qualifiers
source 1..495
/organism="unknown"
BASE COUNT 158 a 96 c 85 g 155 t 1 others
ORIGIN
Query Match 98.0%; Score 79.4; DB 6; Length 495;
Best Local Similarity 98.8%; Pred. No. 4.7e-17;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
Db 126 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 185
QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 186 CAAGAAGGACAGAATCCCAGC 206
RESULT 9
AR272381
LOCUS AR272381 531 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 93 from patent US 6504010.
ACCESSION AR272381
VERSION AR272381.1 GI:29704266
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 531)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W., Mannion,J. and Fan,L.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: US 6504010-A 93 07-JAN-2003;
FEATURES Location/Qualifiers
```

```

source 1..531
/organism="unknown"
BASE COUNT 157 a 139 c 108 g 127 t
ORIGIN
Query Match 98.0%; Score 79.4; DB 6; Length 531;
Best Local Similarity 98.8%; Pred. No. 4.7e-17;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 216
QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 217 CAAGAAGGACAGAATCCCAGC 237
RESULT 10
AR275962
LOCUS AR275962 531 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 93 from patent US 6509448.
ACCESSION AR275962
VERSION AR275962.1 GI:29709607
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 531)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W., Mannion,J., Fan,L. and Wang,A.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: US 6509448-A 93 21-JAN-2003;
FEATURES Location/Qualifiers
source 1..531
/organism="unknown"
BASE COUNT 157 a 139 c 108 g 127 t
ORIGIN
Query Match 98.0%; Score 79.4; DB 6; Length 531;
Best Local Similarity 98.8%; Pred. No. 4.7e-17;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
Db 157 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 216
QY 61 CAAGAAGGACAGAATCCCAGC 81
Db 217 CAAGAAGGACAGAATCCCAGC 237
RESULT 11
AX062466
LOCUS AX062466 531 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 93 from Patent WO0100828.
ACCESSION AX062466
VERSION AX062466.1 GI:12540341
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W. and Mannion,J.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: WO 0100828-A 93 04-JAN-2001;
FEATURES Location/Qualifiers
source 1..531
```


NLPNNNTKPVHDNYVIRRTAVDSGIPLLTNFQVTKLFABAVQSRKVKVDSKSLFPHYRQY
SAGKAA"
BASE COUNT 1484 a 1102 c 1207 g 1422 t
ORIGIN
Query Match 100.0%; Score 81; DB 9; Length 5215;
Best Local Similarity 100.0%; Pred. No. 1e-17;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
Db 4295 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 4354
QY 61 CAAGAAGGACAGAAATCCCAGC 81
Db 4355 CAAGAAGGACAGAAATCCCAGC 4375
RESULT 4
AR184416
LOCUS AR184416 5761 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 1 from patent US 6346382.
ACCESSION AR184416
VERSION AR184416.1 GI:20230381
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5761)
AUTHORS Summar,M.L. and Christman,B.W.
TITLE Human carbamyl phosphate synthetase I polymorphism and diagnostic methods related thereto
JOURNAL Patent: US 6346382-A 1 12-FEB-2002;
FEATURES Location/Qualifiers
source 1. .5761
/organism="unknown"
BASE COUNT 1625 a 1209 c 1309 g 1618 t
ORIGIN
Query Match 100.0%; Score 81; DB 6; Length 5761;
Best Local Similarity 100.0%; Pred. No. 1e-17;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 4359
QY 61 CAAGAAGGACAGAAATCCCAGC 81
Db 4360 CAAGAAGGACAGAAATCCCAGC 4380
RESULT 5
AR184424
LOCUS AR184424 5761 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 11 from patent US 6346382.
ACCESSION AR184424
VERSION AR184424.1 GI:20230389
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5761)
AUTHORS Summar,M.L. and Christman,B.W.
TITLE Human carbamyl phosphate synthetase I polymorphism and diagnostic methods related thereto
JOURNAL Patent: US 6346382-A 11 12-FEB-2002;
FEATURES Location/Qualifiers
source 1. .5761
/organism="unknown"
BASE COUNT 1625 a 1209 c 1309 g 1618 t
ORIGIN
Query Match 100.0%; Score 81; DB 6; Length 5761;
Best Local Similarity 100.0%; Pred. No. 1e-17;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 4359
QY 61 CAAGAAGGACAGAAATCCCAGC 81
Db 4360 CAAGAAGGACAGAAATCCCAGC 4380
RESULT 6
AR184425
LOCUS AR184425 5762 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 13 from patent US 6346382.
ACCESSION AR184425
VERSION AR184425.1 GI:20230390
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5762)
AUTHORS Summar,M.L. and Christman,B.W.
TITLE Human carbamyl phosphate synthetase I polymorphism and diagnostic methods related thereto
JOURNAL Patent: US 6346382-A 13 12-FEB-2002;
FEATURES Location/Qualifiers
source 1. .5762
/organism="unknown"
BASE COUNT 1626 a 1210 c 1308 g 1618 t
ORIGIN
Query Match 100.0%; Score 81; DB 6; Length 5762;
Best Local Similarity 100.0%; Pred. No. 1e-17;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 4359
QY 61 CAAGAAGGACAGAAATCCCAGC 81
Db 4360 CAAGAAGGACAGAAATCCCAGC 4380
RESULT 7
AY166970S36
LOCUS AY166970S36 296 bp DNA linear PRI 28-FEB-2003
DEFINITION Homo sapiens carbamoylphosphate synthetase (CPSI) gene, exon 36; nuclear gene for mitochondrial product.
ACCESSION AY167005
VERSION AY167005.1 GI:28625483
KEYWORDS
SEGMENT 36 of 38
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 296)
AUTHORS Funghini,S., Morrone,A. and Zammarchi,E.
TITLE Human carbamoylphosphate synthetase (CPS I) gene organization
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 296)
AUTHORS Funghini,S., Morrone,A. and Zammarchi,E.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-2002) Dep. of Paediatrics, University of Florence, Via Luca Giordano 13, Florence 50132, Italy
FEATURES Location/Qualifiers
source 1. .296
/organism="Homo sapiens"
/mol_type="genomic DNA"

Query Match 100.0%; Score 81; DB 6; Length 5761;
Best Local Similarity 100.0%; Pred. No. 1e-17;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 4359
QY 61 CAAGAAGGACAGAAATCCCAGC 81
Db 4360 CAAGAAGGACAGAAATCCCAGC 4380
RESULT 6
AR184425
LOCUS AR184425 5762 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 13 from patent US 6346382.
ACCESSION AR184425
VERSION AR184425.1 GI:20230390
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5762)
AUTHORS Summar,M.L. and Christman,B.W.
TITLE Human carbamyl phosphate synthetase I polymorphism and diagnostic methods related thereto
JOURNAL Patent: US 6346382-A 13 12-FEB-2002;
FEATURES Location/Qualifiers
source 1. .5762
/organism="unknown"
BASE COUNT 1626 a 1210 c 1308 g 1618 t
ORIGIN
Query Match 100.0%; Score 81; DB 6; Length 5762;
Best Local Similarity 100.0%; Pred. No. 1e-17;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 60
Db 4300 GCCACATCAGACTGGCTCAACGCCCAACAATGTCCTGCCAACCCAGTGGCATGGCCGTCT 4359
QY 61 CAAGAAGGACAGAAATCCCAGC 81
Db 4360 CAAGAAGGACAGAAATCCCAGC 4380
RESULT 7
AY166970S36
LOCUS AY166970S36 296 bp DNA linear PRI 28-FEB-2003
DEFINITION Homo sapiens carbamoylphosphate synthetase (CPSI) gene, exon 36; nuclear gene for mitochondrial product.
ACCESSION AY167005
VERSION AY167005.1 GI:28625483
KEYWORDS
SEGMENT 36 of 38
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 296)
AUTHORS Funghini,S., Morrone,A. and Zammarchi,E.
TITLE Human carbamoylphosphate synthetase (CPS I) gene organization
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 296)
AUTHORS Funghini,S., Morrone,A. and Zammarchi,E.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-2002) Dep. of Paediatrics, University of Florence, Via Luca Giordano 13, Florence 50132, Italy
FEATURES Location/Qualifiers
source 1. .296
/organism="Homo sapiens"
/mol_type="genomic DNA"

GENE LOGIC INC (US)	
FEATURES	Location/Qualifiers
source	1..5215
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/note="EMBL/GenBank Accession No. D90282"
BASE COUNT	1484 a 1102 c 1207 g 1422 t
ORIGIN	
Query Match	100.0%; Score 81; DB 6; Length 5215;
Best Local Similarity	100.0%; Pred. No. 1e-17;
Matches	81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 GCCACATCAGACTGGCTCAACGCCACACAATGTCCCTGCCAACCCAGTGGCCGTCT 60
Db	4295 GCCACATCAGACTGGCTCAACGCCACACAATGTCCCTGCCAACCCAGTGGCCGTCT 4354
QY	61 CAAGAAGGACAGAATCCCAGC 81
Db	4355 CAAGAAGGACAGAATCCCAGC 4375
RESULT 2	
E04065	
LOCUS	E04065 5215 bp RNA linear PAT 29-SEP-1997
DEFINITION	cDNA encoding human carbamyl phosphate synthetase 1.
ACCESSION	E04065
VERSION	E04065.1 GI:2172275
KEYWORDS	JP 1992335889-A/1.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 5215)
AUTHORS	Haraguchi,Y., Mori,M. and Matsuda,I.
TITLE	CARBAMYLPHOSPHATE SYNTHETASE I GENE, DETECTION OF MUTATION OF THE
	GENE AND DNA PROVE TO USED THEREFOR
JOURNAL	Patent: JP 1992335889-A 1 24-NOV-1992;
	HARAGUCHI YOGO, MORI MASATAKA, MATSUDA ICHIRO
COMMENT	OS Homo sapiens (human)
	PN JP 1992335889-A/1
	PD 24-NOV-1992
	PF 09-MAY-1991 JP 1991135902
	PI HARAGUCHI YOGO, MORI MASATAKA, MATSUDA ICHIRO PC
	C12N15/52,C07H21/04,C12N15/11,C12Q1/68//A61B10/00; CC
	strandedness: Double;
	CC topology: Linear;
	CC hypothetical: No;
	CC anti-sense: No;
	CC *source: tissue_type=Liver;
	FH Key Location/Qualifiers
	FT 5'UTR 1..118
	FT CDS 119..4621
	FT /product='carbamyl phosphate synthetase 1' FT
	3'UTR 4622..5215.
FEATURES	Location/Qualifiers
source	1..5215
	/organism="Homo sapiens"
	/mol_type="genomic RNA"
	/db_xref="taxon:9606"
BASE COUNT	1484 a 1102 c 1207 g 1422 t
ORIGIN	
Query Match	100.0%; Score 81; DB 6; Length 5215;
Best Local Similarity	100.0%; Pred. No. 1e-17;
Matches	81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 GCCACATCAGACTGGCTCAACGCCAACACAATGTCCCTGCCAACCCAGTGGCCGTCT 60
Db	4295 GCCACATCAGACTGGCTCAACGCCAACACAATGTCCCTGCCAACCCAGTGGCCGTCT 4354

QY	61	CAAGAAGGACAGAATCCCAGC 81									
Db	4355	CAAGAAGGACAGAATCCCAGC 4375									
RESULT 3											
HUMCPSI											
LOCUS											
DEFINITION											
HUMCPSI 5215 bp mRNA linear PRI 29-MAY-2002											
Homo sapiens mRNA for carbamyl phosphate synthetase I, complete											
cds.											
D90282											
D90282.1 GI:219552											
KEYWORDS											
SOURCE											
ORGANISM											
Homo sapiens (human)											
Homo sapiens											
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;											
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.											
1 (bases 1 to 5215)											
Haraguchi,Y., Uchino,T., Takiguchi,M., Endo,F., Mori,M. and											
Matsuda,I.											
TITLE											
Cloning and sequence of a cDNA encoding human carbamyl phosphate											
synthetase I: molecular analysis of hyperammonemia											
JOURNAL											
MEDLINE											
PUBMED											
COMMENT											
Gene 107 (2), 335-340 (1991)											
92084128											
1840546											
Submitted (18-JAN-1991) to DDBJ by:											
Yougo Haraguchi											
Department of Pediatrics											
Kumamoto University Medical School											
1-1-1 Honjo											
Kumamoto 860											
Japan											
Phone: 096-344-2111 x5654											
Fax: 096-366-3471.											
Location/Qualifiers											
1..5215											
/organism="Homo sapiens"											
/mol_type="mRNA"											
/db_xref="taxon:9606"											
/chromosome="2"											
/tissue_type="liver"											
/note="short arm"											
119..4621											
/EC_number="6.3.4.16"											
/codon_start=1											
/product="carbamyl phosphate synthetase I"											
/protein_id="BAAL4328.1"											
/db_xref="GI:219553"											
/translation="MTRLTAFKVVTLLKTGFETNVNTAHQKWFSPRGIRLLSVKQAQ											
TAHIVLEDTGTMKGYSGHPSSVAGVVFNTGLGGYPEAITDPAYKGQILTMANPIIG											
NGGAPDTSLSDELGLSKYLESGIKVSGLLVLDYSKDYNHWLATKSLGQWLQEEKVPA											
IYGVDTTMLTKIIRDKGTMLGKIEFEGQPVDFDPNKQNLIAEVSTKDKVKYKGNPT											
KVAVDCGINKNVIRLLVKRGAIEVHLVPWNHDTKMEYDGLIAGPGNPALAEPLIQ											
NVQKILESDRKEPLFGISTGNLITGLAAGAKTYKMSMANRGQNPVLNITNKQAFITA											
QNHCVALDNTLPAGWKPLFVNVDQTNNEGIMHESKPFPAVQFPHPTPGPIDTEYLF											
SFFSLIKKGKATITSVLPKPAIVASRVEVSKVILILSGGLSIGQAGEFDYSGSQAVK											
AMKENVKTVMNPNIASVQINEVGLKQADTVYFLPTPQFVTEVIAEQPDGLILGM											
GGQFALNCGVLEPKRGLKEYGVKVLGTSVESIMATEDRQLFSDKLNINEKIAPSFA											
VESIEDALKAADTIGYPVMIRSAIYALGGLSGGICPNRETLMDSLTKAFAMTNQILVEK											
SVTGKKEIYEYVRDADDNCVTCNMENVDMGVHTGDSVVVAPAQTLNSAEFQMLRR											
TSINVVRLHGIYGEENIQFALHPTSMEYCIIEVNAKMSNSALASKTTGYPLAFTAAR											
IALGFLPGIKNVVSGKTSACPEPSLDYMTKIPRWDLDRFHGTSSRIGSSSMKSVEGV											
MAIGTFEESQFALRMCHPSLEGFTPLPMNKWSPSNLDLRKELSEPSSTRIYAIAR											
AIDNMSLDEIKLTYIDKFLYKMRDILNMEKTLKGLNSESMTETLKRKEIGSDF											
KOISCLGLTEAOTRELRLKKNLHPVVKOIDTLAAEYPSVTNLYVTYNGQEHVDNFD											
DHGMVLCGPHYHIGSVEFDWCVSSIRLRQLRQKKTVVVCNPNETVSTDFDECDKL											
YFEELSLERILDIYHQEACGGCII SVGGQIPNNLAVPLYKNGVKIMTSPLOIDRAED											
RSIFSAYDELKVAQAPWKAVNTLNEALEFAKSDVDPCLLRPSYVLSSGAMNVFSED											
EMKKELEEAATRVSQATPVLLTKVEGAREVEMDVGKGRVISHAISEHVEDAGVHSE											
NATMLPTQOTISOAGIEKVQDATRKIAKAFATSGPFNVFLVKGNDLVNNECNLRASR											
SFPSSKTLGVSTFDVATKVLIGENVDKHLPTLDHPITPVDYVAIKAPMFSWPLRDR											
ADPILRCESMGSTGEVACFGEIGHTAFLKAMLSGTGFKIPQKGLIGIQQSFRPFLGVA											
EQLHNEGFKLFATFATSDWLNANNVPANPVAMPV SQEONPSLSIRKLIRDGSIDLVV											

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 22:25:46 ; Search time 1416.5 Seconds
(without alignments)
2339.343 Million cell updates/sec

Title: US-09-585-077C-1_COPY_4300_4380
Perfect score: 81
Sequence: 1 gccacatcagactggtctaa.....aagaaggacagaatcccagc 81

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	81	100.0	5215	6	AX409122	AX409122 Sequence
2	81	100.0	5215	6	E04065	E04065 cDNA encodi
3	81	100.0	5215	9	HUMCPSI	D90282 Homo sapien
4	81	100.0	5761	6	ARI84416	ARI84416 Sequence
5	81	100.0	5761	6	ARI84424	ARI84424 Sequence
6	81	100.0	5762	6	ARI84425	ARI84425 Sequence
7	79.4	98.0	296	9	AY166970S36	AY167005 Homo sapi
8	79.4	98.0	495	6	ARI84418	ARI84418 Sequence
9	79.4	98.0	531	6	AR272381	AR272381 Sequence
10	79.4	98.0	531	6	AR275962	AR275962 Sequence
11	79.4	98.0	531	6	AX062466	AX062466 Sequence
12	79.4	98.0	531	6	AX367383	AX367383 Sequence
13	79.4	98.0	4503	9	HSCPSIMUT	Y15793 Homo sapien
14	79.4	98.0	5761	6	ARI84417	ARI84417 Sequence
15	79.4	98.0	5761	9	AF154830	AF154830 Homo sapi
16	79.4	98.0	121959	9	AF536523	AF536523 Homo sapi
17	79.4	98.0	183696	9	AC008172	AC008172 Homo sapi
18	73	90.1	133	10	RATCPSIX36	Z27487 R.norvegicu
19	73	90.1	147	10	RATCPSI11	M12327 Rattus norv
20	73	90.1	229369	2	AC106119	AC106119 Rattus no
21	69.8	86.2	255581	2	AC101854	AC101854 Mus muscu
22	69	85.2	816	6	BD091111	BD091111 DNA chip
23	47	58.0	4716	5	RCU05193	U05193 Rana catesb
24	46	56.8	580	6	AX525511	AX525511 Sequence
25	41	50.6	4597	5	OMU65893	U65893 Oncohyncus
26	38.4	47.4	4717	5	AF006491	AF006491 Micropter
27	33	40.7	63058	2	AC133885	AC133885 Homo sapi
28	30.6	37.8	31	6	AX249689	AX249689 Sequence
29	30	37.0	6001	5	SQUCARPSYN	L31362 Squalus aca
30	29.6	36.5	4755	5	AF119250	AF119250 Oreochrom
31	29.6	36.5	222983	2	AC129695	AC129695 Rattus no
32	29.6	36.5	253973	2	AC094544	AC094544 Rattus no
33	29.6	36.5	283158	2	AC137193	AC137193 Rattus no
34	29.6	36.5	320205	2	AC106178	AC106178 Rattus no
35	29.4	36.3	36307	9	HS366D1	Z97986 Human DNA s
36	29.4	36.3	256073	9	AE006464	AE006464 Homo sapi
37	29.2	36.0	154736	9	AL139246	AL139246 Human DNA
38	29	35.8	1101	11	G67204	G67204 Xq4133 KWOK
39	29	35.8	141672	9	HS142F18	AL031073 Human DNA
40	28	34.6	63058	2	AC133885	AC133885 Homo sapi
41	28	34.6	72620	2	AC121127	AC121127 Mus muscu
42	28	34.6	161065	9	AC124799	AC124799 Homo sapi
43	28	34.6	161190	9	AC005137	AC005137 Human Chr
44	28	34.6	234487	2	AC094748	AC094748 Rattus no
45	27.8	34.3	5989	5	AF169248	AF169248 Opsanus b

ALIGNMENTS

RESULT 1	AX409122	Sequence 1769 from Patent WO0229103.	5215 bp	DNA	linear	PAT 14-JUN-2002
AX409122	LOCUS	AX409122				
DEFINITION	Sequence 1769 from Patent WO0229103.					
ACCESSION	AX409122					
VERSION	AX409122.1	GI:21441827				
KEYWORDS						
SOURCE						
ORGANISM						
	Homo sapiens (human)					
	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1					
AUTHORS	Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.					
TITLE	Gene expression profiles in liver cancer					
JOURNAL	Patent: WO 0229103-A 1769 11-APR-2002;					